

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 23, 2004, 07:50:32 ; Search time 8804 Seconds
(without alignments)
11124.144 Million cell updates/sec

Title: US-10-085-198-47
Perfect score: 2071
Sequence: 1 cccccgcgccccgcgcg.....ctatcagtcctgtcttc 2071

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|--------------------|
| 1 | 2071 | 100.0 | 2071 | 6 | AX921054 Sequence |
| 2 | 1634.8 | 78.9 | 3457 | 9 | BC041772 Homo sapi |
| 3 | 1525.8 | 73.7 | 3179 | 6 | AX179748 Sequence |
| 4 | 1349.8 | 65.2 | 3273 | 10 | BC011467 Mus muscu |
| 5 | 1161.4 | 56.1 | 1353 | 6 | CQ725703 Sequence |
| 6 | 692.2 | 33.4 | 2342 | 6 | BD242859 Secreted |
| 7 | 563.8 | 27.2 | 1587 | 6 | AX481748 Sequence |
| 8 | 563.8 | 27.2 | 1836 | 9 | BC006156 Homo sapi |
| 9 | 563.8 | 27.2 | 2017 | 9 | BC038961 Homo sapi |
| 10 | 563.8 | 27.2 | 2049 | 9 | BC065235 Homo sapi |
| 11 | 563.8 | 27.2 | 2050 | 9 | BC008325 Homo sapi |
| 12 | 563.8 | 27.2 | 2165 | 9 | AF212371 Homo sapi |
| 13 | 563.8 | 27.2 | 2230 | 6 | AX481746 Sequence |
| 14 | 542.8 | 26.2 | 2201 | 6 | BD191275 186 human |
| 15 | 542.8 | 26.2 | 2201 | 6 | AX924785 Sequence |
| 16 | 532 | 25.7 | 2171 | 10 | BC026854 Mus muscu |
| 17 | 530.4 | 25.6 | 1731 | 10 | BC002297 Mus muscu |
| 18 | 524 | 25.3 | 2211 | 10 | AF212372 Mus muscu |
| 19 | 512.4 | 24.7 | 1661 | 6 | BD191276 186 human |

| | | | | | |
|----|-------|------|--------|---|--------------------|
| 20 | 512.4 | 24.7 | 1661 | 6 | AX924786 Sequence |
| 21 | 453 | 21.9 | 1460 | 6 | BD191122 186 human |
| 22 | 453 | 21.9 | 1460 | 6 | AX924632 Sequence |
| 23 | 452.4 | 21.8 | 2614 | 5 | BC048024 Danio rer |
| 24 | 449.2 | 21.7 | 2643 | 5 | AF465772 |
| 25 | 437.8 | 21.1 | 1915 | 9 | AK172832 Homo sapi |
| 26 | 437 | 21.1 | 112130 | 2 | AC135621 |
| 27 | 437 | 21.1 | 165762 | 2 | AC124263 Homo sapi |
| 28 | 437 | 21.1 | 178714 | 2 | AC129504 Homo sapi |
| 29 | 437 | 21.1 | 185378 | 9 | AC118754 Homo sapi |
| 30 | 437 | 21.1 | 190466 | 2 | AC022285 Homo sapi |
| 31 | 437 | 21.1 | 190466 | 2 | AC129497 Homo sapi |
| 32 | 434.6 | 21.0 | 1539 | 6 | AX328008 Sequence |
| 33 | 434.6 | 21.0 | 1897 | 6 | AX709979 Sequence |
| 34 | 434.6 | 21.0 | 2121 | 6 | AX328006 Sequence |
| 35 | 428.8 | 20.7 | 154622 | 2 | AC126336 Homo sapi |
| 36 | 399 | 19.3 | 56981 | 9 | AC126391 Homo sapi |
| 37 | 399 | 19.3 | 185762 | 2 | AC124263 |
| 38 | 399 | 19.3 | 168070 | 2 | AC126328 Homo sapi |
| 39 | 399 | 19.3 | 178714 | 2 | AC129504 |
| 40 | 399 | 19.3 | 180317 | 2 | AC025590 Homo sapi |
| 41 | 399 | 19.3 | 185378 | 9 | AC118754 Homo sapi |
| 42 | 399 | 19.3 | 190466 | 2 | AC022285 Homo sapi |
| 43 | 399 | 19.3 | 199463 | 2 | AC129497 Homo sapi |
| 44 | 388 | 18.7 | 35758 | 2 | AC127539 Homo sapi |
| 45 | 368 | 17.8 | 1431 | 6 | AX191614 Sequence |

ALIGNMENTS

RESULT 1
AX921054
LOCUS AX921054 2071 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 47 from Patent WO02068652.
ACCESSION AX921054
VERSION AX921054.1 GI:40214759
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Nov-x proteins and nucleic acids encoding same
TITLE Patent: WO 02068652-A 47 06-SEP-2002;
JOURNAL Location/Qualifiers
FEATURES
source
1..2071
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2071; DB 6; Length 2071;
Best Local Similarity 100.0%; Pred. No. 1.7e-290;
Matches 2071; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCCCCCGCCCGATCCCGCGCGCATGATGCTGGAATCGCGCTCGCGCGCGCG 60
Db 1 CCCCCCGCCCGATCCCGCGCGCATGATGCTGGAATCGCGCTCGCGCGCGCG 60
Qy 61 GCGCGCGCGAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 GCGCGCGCGAGGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy 121 CGAGGGGCTCGCGGTAGCGTTCGTCGGGGCGCGGGCGCGCGCGCGCGCGT 180
Db 121 CGAGGGGCTCGCGGTAGCGTTCGTCGGGGCGCGGGCGCGCGCGCGCGT 180
Qy 181 GCGCGCGCGATGAGGTGCGAGCGTTCGTCGGGCGCGGTAAAGCGGGCCCC 240
Db 181 GCGCGCGCGATGAGGTGCGAGCGTTCGTCGGGCGCGGTAAAGCGGGCCCC 240

| | | | |
|----|------|--|------|
| QY | 241 | CCGGGACCCCGGCACCCCGGCTGGCAGCTACTGCAAAAGGGCCCCGGCGCTCAGCAG | 300 |
| DB | 241 | CCGGGACCCCGGCACCCCGGCTGGCAGCTACTGCAAAAGGGCCCCGGCGCTCAGCAG | 300 |
| QY | 301 | CCCAAAACCGGCACGCTTGGGCGCGGGCGGGGGGAGCGCGCCACTCCTCAGCTTGGGC | 360 |
| DB | 301 | CCCAAAACCGGCACGCTTGGGCGCGGGCGGGGGGAGCGCGCCACTCCTCAGCTTGGGC | 360 |
| QY | 361 | AACGTGCTCAACTACTCTGAGCAAGGTACACGCTGGCAGGCGTCCTTCTGGACATCCAGCAG | 420 |
| DB | 361 | AACGTGCTCAACTACTCTGAGCAAGGTACACGCTGGCAGGCGTCCTTCTGGACATCCAGCAG | 420 |
| QY | 421 | CACCTTGGGGTCAAGGACCGAGCGCGCGGCTCTGCGATCAGTGTTCATCTGTAGCTTC | 480 |
| DB | 421 | CACCTTGGGGTCAAGGACCGAGCGCGCGGCTCTGCGATCAGTGTTCATCTGTAGCTTC | 480 |
| QY | 481 | ATGGTGGCTGCCGCCATCTTCGGCTACCTTGGGCGACCGCTTCAACAGGAAGGTGATTCTC | 540 |
| DB | 481 | ATGGTGGCTGCCGCCATCTTCGGCTACCTTGGGCGACCGCTTCAACAGGAAGGTGATTCTC | 540 |
| QY | 541 | AGTGGCGGATTTCTTCTGGTGGCGGCTCACTTCTCCAGTCTCTTCATTCGCCAGCAG | 600 |
| DB | 541 | AGTGGCGGATTTCTTCTGGTGGCGGCTCACTTCTCCAGTCTCTTCATTCGCCAGCAG | 600 |
| QY | 601 | TACTTCTGGGTGCTGGTCTCTGTCGGGGGCTGGTGGGCATCGGGGAGCCAGCTACTCC | 660 |
| DB | 601 | TACTTCTGGGTGCTGGTCTCTGTCGGGGGCTGGTGGGCATCGGGGAGCCAGCTACTCC | 660 |
| QY | 661 | ACCATGCCGCCCACTATCATTTGGCGACCTTCTCACCAAGAACACGGTACGTCTAGCTG | 720 |
| DB | 661 | ACCATGCCGCCCACTATCATTTGGCGACCTTCTCACCAAGAACACGGTACGTCTAGCTG | 720 |
| QY | 721 | TGGTCTTCTACTTTCGCCATCCACTGGGCGAGTGGCTGGGCTACATTACTGGCTCCAGC | 780 |
| DB | 721 | TGGTCTTCTACTTTCGCCATCCACTGGGCGAGTGGCTGGGCTACATTACTGGCTCCAGC | 780 |
| QY | 781 | GTGAAGCAGCAGCGCGGAGACTGGCACTGGGCAATGGGGTGTCCCTGTCTCTGGGCATG | 840 |
| DB | 781 | GTGAAGCAGCAGCGCGGAGACTGGCACTGGGCAATGGGGTGTCCCTGTCTCTGGGCATG | 840 |
| QY | 841 | ATCACAGGAACACTCATCTCTCATTTGGTCCGAGCCACTAAAGAGGGGTATGCCGACCAG | 900 |
| DB | 841 | ATCACAGGAACACTCATCTCTCATTTGGTCCGAGCCACTAAAGAGGGGTATGCCGACCAG | 900 |
| QY | 901 | CTGGGACACAGCTCAAGGCGCGGACTCATGCTCCGATATGAAGGCCCTGATTCTCGA | 960 |
| DB | 901 | CTGGGACACAGCTCAAGGCGCGGACTCATGCTCCGATATGAAGGCCCTGATTCTCGA | 960 |
| QY | 961 | AACCGCAGCTACGTCTTCTCTCTCCCTGGCCACGTGGCTCTCTCTTCGCCACGGGGGCC | 1020 |
| DB | 961 | AACCGCAGCTACGTCTTCTCTCTCCCTGGCCACGTGGCTCTCTCTTCGCCACGGGGGCC | 1020 |
| QY | 1021 | CTGGGCAATGTGGATCCCGCTCTACTCTGCA CGCGGCCCAAGTTGTGCAAGAGACAGCAG | 1080 |
| DB | 1021 | CTGGGCAATGTGGATCCCGCTCTACTCTGCA CGCGGCCCAAGTTGTGCAAGAGACAGCAG | 1080 |
| QY | 1081 | ACGTGCAACACGCCCTGTGGGGCCAAAGGACAGCTCATCTTTGGGGCCATCACCTGCG | 1140 |
| DB | 1081 | ACGTGCAACACGCCCTGTGGGGCCAAAGGACAGCTCATCTTTGGGGCCATCACCTGCG | 1140 |
| QY | 1141 | TTTACGGGATTTCTGGGCGTGGTCA CGGGGGCAGGACCGCGCTGTGTGCGCGCTGAAG | 1200 |
| DB | 1141 | TTTACGGGATTTCTGGGCGTGGTCA CGGGGGCAGGACCGCGCTGTGTGCGCGCTGAAG | 1200 |
| QY | 1201 | ACCGAGCGGGCGGACCCACTGGTGTGTGCGTGGGCATGCTGGGCTCTGCCATCTTCATC | 1260 |
| DB | 1201 | ACCGAGCGGGCGGACCCACTGGTGTGTGCGTGGGCATGCTGGGCTCTGCCATCTTCATC | 1260 |
| QY | 1261 | TGCCTGATCTTCGTGGCTGCCAAGACAGCAGCATCGTAGGAGCCTATATCTGTATCTCGTC | 1320 |
| DB | 1261 | TGCCTGATCTTCGTGGCTGCCAAGACAGCAGCATCGTAGGAGCCTATATCTGTATCTCGTC | 1320 |
| QY | 1321 | GGGGAGACGCTGCTGTTTCTTCACTGGGCGCATCACTGCAGACATCTCTCATGTACGTGGTC | 1380 |

| | | | | | |
|----|------|--|--|------------------------------------|------|
| Db | 1321 | | GGGAGACGCTGCTGTTTCTTAAC | TGGGCATCACTGCAGACATCTCATGTACGTGGTC | 1380 |
| Qy | 1381 | | ATCCCAACGGGGCGCGCACTGCGTGGCTTGACAGACTTACCTCCCACTGCTGGGG | | 1440 |
| Db | 1381 | | ATCCCAACGGGGCGCGCCATGCGGTGGCTTGACAGACTTACCTCCCACTGCTGGGG | | 1440 |
| Qy | 1441 | | GAGCCGGGAGCCCTCACTCATTTGGCTTTATCTCAGACCTGATCCGCAGACACTAAG | | 1500 |
| Db | 1441 | | GAGCCGGGAGCCCTACCTCATTTGGCTTTATCTCAGACCTGATCCGCAGACACTAAG | | 1500 |
| Qy | 1501 | | GATCCCGGCTTGAGGATTCCTGAGCCCTGGGCTACGCGTCACTGCTGCCCTTTGCTC | | 1560 |
| Db | 1501 | | GACTCCCGGCTCTGGGAGTTCCCTGAGCCCTGGGCTACGCGTCACTGCTGCCCTTTGCTC | | 1560 |
| Qy | 1561 | | GTGTCCTCTGGCGGCATGTTCTTCTCGCCACATGCGCTCTTCTTCTGTCAGCAGCCGGCC | | 1620 |
| Db | 1561 | | GTGTCCTCTGGCGGCATGTTCTTCTCGCCACATGCGCTCTTCTTCTGTCAGCAGCCGGCC | | 1620 |
| Qy | 1621 | | AGGCTGAGCAGACACTTGGGGGAGACGGCGGGGTCAAGGTGGTGCATCAGCGGGG | | 1680 |
| Db | 1621 | | AGGCTGAGCAGACACTTGGGGGAGACGGCGGGGTCAAGGTGGTGCATCAGCGGGG | | 1680 |
| Qy | 1681 | | CCGGGCCC GGSCACTGCTCTGGCACA TCGTGTCTGTGGGGCCAGCTGACCGGAGTCTG | | 1740 |
| Db | 1681 | | CCGGGCCC GGSCACTGCTCTGGCACA TCGTGTCTGTGGGGCCAGCTGACCGGAGTCTG | | 1740 |
| Qy | 1741 | | GGCAGGAGCTCTGTACGCCACAGGGGAGATGGGAGAGCCACGGGTGGGGAGAGAGAG | | 1800 |
| Db | 1741 | | GGCAGGAGCTCTGTGTAGCCACAGGGGAGATGGGAGAGCCACGGGTGGGGAGAGAGAG | | 1800 |
| Qy | 1801 | | AGAGGNGTAAAGAGGAAAGGAAAGAGTCAAGAAAGTAAAGAGAAAGGGAGGGGCC | | 1860 |
| Db | 1801 | | AGAGGAGAGTAAAGAGGAAAGGAAAGAGTCAAGAAAGTAAAGAGAAAGGGAGGGGCC | | 1860 |
| Qy | 1861 | | CAGCTTTGAAACCACTAAGTCCAGAGACAAACCCAAAGTCTGGATCCACAGACACCCC | | 1920 |
| Db | 1861 | | CAGCTTTGAAACCACTAAGTCCAGAGACAAACCCAAAGTCTGGATCCACAGACACCCC | | 1920 |
| Qy | 1921 | | GTGGCTCCCA CAGCTCCAGGCTGACCTGGCACTGGGCGCTCAGGGCTGAGCCCCAGCAA | | 1980 |
| Db | 1921 | | GTGGCTCCCA CAGCTCCAGGCTGACCTGGCACTGGGCGCTCAGGGCTGAGCCCCAGCAA | | 1980 |
| Qy | 1981 | | CCAGTGGGTGCACTGAGTGCATGGGAGGTCTGTACCTTCCCGGCCCA CCCCAGGGCAGG | | 2040 |
| Db | 1981 | | CCAGTGGGTGCACTGAGTGCATGGGAGGTCTGTACCTTCCCGGCCCA CCCCAGGGCAGG | | 2040 |
| Qy | 2041 | | GCTACGCTGGCTATCAGGTCCTCTGCTTCC | 2071 | |
| Db | 2041 | | GCTACGCTGGCTATCAGGTCCTCTGCTTCC | 2071 | |

RESIT.T 2

RESOL 2
RC041772

LOCUS

DEFINITION

1
2
3
4
5
6
7
8

ACCESSIO

VERSION

KEYWORDS

SOURCE

ORGANI

1

REFERENCE

AUTHOR
ETTY D.

STILL


```
Qy 1586 TCGCCACTGCGCTCTTCTTCTGAGCGACCGCGCCAGGGCTGAGCAGCACCTG 1638
Db 1501 TCGCCACTGCGCTCTTCTTCTGAGCGACCGCGCCAGGGCTGAGCAGCAGTG 1553

RESULT 4
BC011467 3273 bp mRNA linear ROD 16-SEP-2003
LOCUS Mus musculus cDNA sequence BC011467, mRNA (cDNA clone
DEFINITION IMAGE:3154539), partial cds.
ACCESSION BC011467
VERSION BC011467.1 GI:15079261
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3273)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Narusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, K.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, R.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2238257
MEDLINE 12477932
PUBMED 2 (bases 1 to 3273)
REFERENCE Strausberg, R.
Direct Submission
Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: Lothar Hengnighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILML at: http://image.llnl.gov
Series: IRAC Plate: 23 Row: b Column: 10.
Location/Qualifiers
1. .3273
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"

FEATURES
source
```

gene

```
/clone="IMAGE:3154539"
/tissue_type="Mammary tumor, Brcal-1; MMTV-Cre model. 10
months old, gross tissue."
/clone_lib="NCI CGAP_Mam3"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
<1. .3273
/gene="BC011467"
/notes="synonym: MGC37865"
/db_xref="LocusID:216892"
/db_xref="MGI:2384936"
<1. .1775
/gene="BC011467"
/codon_start=3
/product="BC011467 protein"
/protein_id="AAH11467.1"
/db_xref="GI:15079262"
/db_xref="LocusID:216892"
/db_xref="MGI:2384936"
/translators="DGFGRDGAAPRAAHALSGPLGLASPEPPAPPAAPARAGMM
LECAAAAGAGAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAEAE
GVRVPSGLPSIPSTPGCAAAKSPAPQKPSLGRGAAAIISLGNLVNLDL
YTAVLLDIQHPGVKDRGAGLLQSVFICSPWAAPIFGYLGDRFRKXVILSGIFF
WSAVTSSSFIPOQYFWLLVLSGLVGEASYSYTIAPTIIGDUFTKNTKRLMUSVY
FAIPLGSGLYITGSSVQAAGDMHVALRVSPVLMITGTLILVLPATKRGHADQLG
GOLKARTSWLRDMKALIRNRYVFSLSATSAVSATGALGMWIPLYLHRAQVQVGTAE
TCNSPCGAKSLIFGATCTGTGLVVTGATRCWLRCATRADPLVCAVQVQKGTAE
FICLIFAAKTSIVGAVICIFVGETLLFSNNAITADILMYVVIPTREATAVALQSPFS
HLIGDAGSPYLIIGFISDLIRQSKDSPLNEFLSLGYALMLCPFFVVLVGGMFFLATALF
FLSDRAAEQQVNLVNPVPSVKV"
549. .914
/misc_feature
/gene="BC011467"
/notes="sugar tr; Region: Sugar (and other) transporter"
/db_xref="CDB:pfam00083"
```

ORIGIN

```
Query Match 65.2%; Score 1349.8; DB 10; Length 3273;
Best Local Similarity 85.6%; Pred. No. 4.3e-186;
Matches 1501; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

Qy 1 CCCCCCGCCCGCCGATCCGCGCGCGGCGGATGATGCTGGAATCGGCTCGCGCGCGCG 60
Db 99 CCCCCCGCCCGCCCGCGCGCGCGCGCGGCGGATGATGCTGGAATCGGCTCGCGCGCGCG 158

Qy 61 GCGCGCGCGCGAGAGAGAGAGAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 159 GCGCGCGCGCGAGAGAGAGAGAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 218

Qy 121 CGAGGGGCTCGCGGTAGCGGTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 180
Db 219 CCAGGGGCTGAGGTAGCGGTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 278

Qy 181 GCGCGCGCGCGATGAGGTGCGAGCGGTGTCGGGCGAGCGGTAAAGCGGGGCGCGCGCGG 240
Db 279 TCGCGCGCGCGAGAGAGGTGCGAGCGGTGTCGGGCGAGCGGTGAGCGGGGCGCGCGCGG 338

Qy 241 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 339 CCAGCGATCCCCAGCACCCCTTGGCTGCGCAGCAGCTGCCAAGGGGCGCGCGCGCGCG 398

Qy 301 CCCAAACCGCGCGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 399 CCCAAACCGCGCGAGCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 458

Qy 361 AACGTGCTCAACTACTGAGACAGGTACACCGTGGCAGGCGCTCTTGTGGACATCCAGCAG 420
Db 459 AACGTGCTCAACTACTGAGACAGGTACACGTGGCAGGCTGCTTCTTGGACATCCAGCAG 518

Qy 421 CACTTTGGGTCAGGACCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db 519 CACTTTGGGTCAGGACCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 578

Qy 481 ATGGTGGCTGCCCCCATCTTTCGGCTACCTGGGCGACCGCTTCAACAGGAAGGTGATTCTC 540
```

Db 579 ATGGTGGCTGCCCCCATCTTTGGCTACTCTGGGAGCCGCTTCAATAGAAAGGTGATCTC 638
Qy 541 AGCTGGCGATTTCTTCTGGTGGCGGTCACCTTTCTCCAGCTCTTCAATCCCGAGAG 600
Db 639 AGCTGGCGATCTTCTTCTGGTCTGTGTACCTTCTCGAGTTCTTCAATCCCTCAGCAG 698
Qy 601 TACTTCTGGCTGCTGTCTGTCCGGGGCTGGTGGGATCGGGAGGCGAGCTACTCC 660
Db 699 TACTTCTGGCTACTAGTCTGTCCGAGGGCTAGTAGGTATTTGGTGAAGCCAGCTACTCC 758
Qy 661 ACCATCGCCCCCACTATATCTGCGACCTTTCCACCAAGAACACGCGTACGCTCATGCTG 720
Db 759 ACCATAGACCCACCATCATTTGGGACCTTTGTTTACCAAGAACACGCGCATTTGATGCTA 818
Qy 721 TCCGCTCTTCTACTTGGCCATCCCACTGGGCACTGGGCTGAGTGGGCTACATTAATCGCTCCAGC 780
Db 819 TCTGTCTTCTATTTTGGCCATCCCCCTGGGCACTGGGCTGAGTATATACAGGTTCCAGC 878
Qy 781 GTGAAGCAGGACCGCGAGACTGGGCACTGGGCAATTTGGGGGTGTCCCTGTCTTGGGCATG 840
Db 879 GTGAAGCAGGACGCTGGGCACTGGGCAATTTGGGGGTGTCCCTGTCTTGGGCATG 938
Qy 841 ATCAGAGGAACACTCATCTCTATCTTGTGTCAGGCACTAAAGGGGTGATCCGACACAG 900
Db 939 ATCAGAGGAACACTCATCTCTATCTTGTGTCAGGCACTAAAGGGGTGATCCGATCA 998
Qy 901 CTCGGGACAGCTCAAGGCCCGGACCTCATGGCTCCGAGATATGAAGCCCTGATTCGA 960
Db 999 CTTGGGGGAGCTCAAGCAGGACCTCTGGCTTCGAGACATGAAGCCCTGATTCGA 1058
Qy 961 AACCGCAGTACGTCTTCTCTCTGGCCAGCTGGGCTGTCTCTTCCGACGGGGCC 1020
Db 1059 AACCGCAGTACGTCTTCTCTCTGGCCAGCTGGGCTGTCTCTTCCGACGGGGCC 1118
Qy 1021 CTGGGCACTGGATCCGCTCTACCTGTCACCGCGCCAAAGTTGTGACAGAGCAGCAGAG 1080
Db 1119 CTGGGCACTGGATCCGCTCTATCTTCTACCGTGTCTCAAGTTGTGACAGAGCAGCAGAG 1178
Qy 1081 ACGTGCAACAGCCCGCCCTGTGGGGCCAAAGCAGCCTCATCTTTGGGGCCATCACTGTC 1140
Db 1179 ACTTGCACAGCCCGCCCTGTGGAGCCAAAGCAGCCTCATCTTTGGAGCCATTAACCTGC 1238
Qy 1141 TTTACGGGATTTCTGGGCTGTGTCAAGGGGCAAGAGCCACCGGCTGGTGGCCCTGAAG 1200
Db 1239 TTTACTGGCTTCTGGGCTGTGTCAAGGGGCAAGAGCCACCTGCTTGGTGGCCCTGGGG 1298
Qy 1201 ACCCAGGGGCGACCCCACTGGTGTGCGGTCGCGGCACTGCGGCTCGCCATCTTCATC 1260
Db 1299 ACTCAACGTGCTGACCCATTTGGTGTGCTGTGGGCACTGCGGATCCGCGCATCTTCATC 1358
Qy 1261 TGCTGATCTTCTGGCTGCGCAAGAGCAGCATCGTAGGAGCCCTATATCTGTATCTTCGTC 1320
Db 1359 TGCTCATCTTTGTGGCTGCCAAGACAGCATTTGTAGGCGCTTATATCTGCACTTTTGT 1418
Qy 1321 GGGGAGACGCTGTGTTTCTAACTGGGCACTACCTGCGAGACATCTCATGTATGATGCTGTC 1380
Db 1419 GGAGAGACACTGCTGTTTCTTACTGGGCACTACCTGCGAGACATCTCATGTATGATGCTGTC 1478
Qy 1381 ATCCCAAGGGGCGGCGCACTGCGGTCGCGGCTTGCAGAGCTTCACTCCGACCTGCTGGGG 1440
Db 1479 ATCCCAAGGGGCGGCGCACTGCGGTCGCGGCTTGCAGAGCTTCACTCCGACCTGCTGGGG 1538
Qy 1441 GAGCGGGGAGCCCTTACCTCATTTGGCTTTTATCTCAGACCTGATCCGCGCAGAGCACTAAG 1500
Db 1539 GAGCTTGGAGCCCTTACCTCATTTGGCTTTATCTCGAACCTTATCCGCGCAGAGCACTAAG 1598
Qy 1501 GACTCCCGCTCTGGAGGTTCTTGAGCCTGGGCTACGCGCTCATGTCTGCTTCCCTTTCGTC 1560
Db 1599 GACTCCCGCTCTGGAGGTTCTTGAGCCTGGGCTATGCGCTCATGTCTGCTTCCCTTTCGTC 1658
Qy 1561 GTGGTCTGGGGGCACTGCTTCTCTCGCACCTGCGCTCTTCTTCTGTCAGCGACCGCGCC 1620

Db 1659 GTGGTCTGGGTGGCATGTTCTTCTTCCACTGCTCTCTTCTTCTTCTCAGCGACCGTGC 1718
Qy 1621 AGGCTGAGCAGCACCTGGGGAGAGAGCGGGGTTCAGGTTGGTGCATCAGCGGGG 1680
Db 1719 AAGCTGAGCAGCAGGTGAACAGCTGGTGTATGCTCTCGATCCTGTAAGTCTGAGT 1778
Qy 1681 CCGGCCCGGCACTGCTTCTGGCAGCATGCTGTCTGGGGGGCAGCTGACCGAGGTGCTG 1740
Db 1779 GGTGGCTGGGACAGCAATGAATGCACTGAGTTAGTTGGGAGAGGTTCTCCAGCATCTCG 1838
Qy 1741 GCGAGGACCTCG 1753
Db 1839 GACCGCAGGACG 1851

RESULT 5
LOCUS CQ725703 1353 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 11637 from Patent WO02068579.
ACCESSION CQ725703
VERSION CQ725703.1 GI:42287032
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 11637 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..1353
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 56.1%; Score 1161.4; DB 6; Length 1353;
Best Local Similarity 99.5%; Pred. No. 9.5e-159;
Matches 1165; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 468 CATCTGTAGCTTCATGTTGGCTGCCGCCCATCTTCGGCTACCTGGGCGCGCTTCAACAG 527
Db 144 CGTTGGAAGCTTCATGGTGGCTGCCGCCCATCTTCGGCTACCTGGGCGCGCTTCAACAG 203
Qy 528 GAAGGTGATTTCTCAGCTGGGCAATTTTCTTCTGGTGGGCGGTACACCTTCTCCAGCTTCTT 587
Db 204 GAAGGTGATTTCTCAGCTGGGCAATTTTCTTCTGGTGGGCGGTACACCTTCTCCAGCTTCTT 263
Qy 588 CATTTCCCGCAGCATTTCTTGGCTGCTGGTCTCTGTCCTCCGGGGCTGGTGGGCAATCGGGGA 647
Db 264 CATTTCCCGCAGCATTTCTTGGCTGCTGGTCTCTGTCCTCCGGGGCTGGTGGGCAATCGGGGA 323
Qy 648 GCGCAGCTACTCCACCATCGCCCCCACCATCATATTGGCGACCTTTCACCAAGAACACGCG 707
Db 324 GCGCAGCTACTCCACCATCGCCCCCACCATCATATTGGCGACCTTTCACCAAGAACACGCG 383
Qy 708 TAGCGTCACTGCTGCTCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 767
Db 384 TAGCGTCACTGCTGCTCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 443
Qy 768 TACTGGCTCCAGCTGAAGCAGCGAGCTGGCACTGGCACTGGGCAATTCGGGCTGTCGCC 827
Db 444 TACTGGCTCCAGCTGAAGCAGCGAGCTGGCACTGGCACTGGGCAATTCGGGCTGTCGCC 503
Qy 828 TGTCTTGGGCACTGATCAGGAACTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 887
Db 504 TGTCTTGGGCACTGATCAGGAACTCATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 563
Qy 888 TCATGCGCAGCTCGGGGACAGCTCAAGGCGCGGACCTCATGCGCTCCGAGATATGAA 947

Db 564 TCATGCCACACAGCTCGGGGACACAGCTCAAGGCCGAGACCTATGGCTCCGAGATATGAA 623
Qy 948 GSCCTGATTTGAAACCCAGCTAGCTTCTCTCCCTGCGCCAGCTGCGGTGCTCTCTT 1007
Db 624 GSCCTGATTCGAAACCCAGCTAGCTTCTCTCCCTGCGCCAGCTGCGGTGCTCTCTT 683
Qy 1008 CGCACCGGGGCTTGGGCTGATGATCCGCTGCTACCTGCAACCGCGCCCAAGTTGTGCA 1067
Db 684 CGCACCGGGGCTTGGGCTGATGATCCGCTGCTACCTGCAACCGCGCCCAAGTTGTGCA 743
Qy 1068 GAAGACAGCAGAGCTCAACAGCCGCTGTTGGGCGCAAGACAGCAGCTCATCTTTGG 1127
Db 744 GAAGACAGCAGAGCTCAACAGCCGCTGTTGGGCGCAAGACAGCAGCTCATCTTTGG 803
Qy 1128 GGCATCAGCTGTTTACGGGATTTCTGGGCTGCTGCTCAAGCGGCGAGGACCGCGCTG 1187
Db 804 GGCATCAGCTGTTTACGGGATTTCTGGGCTGCTGCTCAAGCGGCGAGGACCGCGCTG 863
Qy 1188 GTGCCGCTGAAGACCCAGCGGGCGACCACTGCTGCTGCTGCGGCTGCTGAGGCTATAT 1247
Db 864 GTGCCGCTGAAGACCCAGCGGGCGACCACTGCTGCTGCTGCGGCTGCTGAGGCTATAT 923
Qy 1248 TGCCATCTTATCTGCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1307
Db 924 TGCCATCTTATCTGCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
Qy 1308 CTGTATCTTCTGCGGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1367
Db 984 CTGTATCTTCTGCGGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
Qy 1368 CATGTACGTGCTATCCCAACCGCGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1427
Db 1044 CATGTACGTGCTATCCCAACCGCGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1103
Qy 1428 CCACCTGCTGGGAGCGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1487
Db 1104 CCACCTGCTGGGAGCGCGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163
Qy 1488 CCAGACACTAAGACTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1547
Db 1164 CCAGACACTAAGACTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1223
Qy 1548 CTGCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1607
Db 1224 CTGCCCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1283
Qy 1508 CAGGACCGCGCAGGGCTGAGCAGCCTG 1638
Db 1284 CAGGACCGCGCAGGGCTGAGCAGCAGGTG 1314

RESULT 6

BD242859 2342 bp DNA linear PAT 17-JUL-2003
LOCUS Secreted proteins and polynucleotides encoding them.
DEFINITION
ACCESSION BD242859
VERSION BD242859.1 GI:33052629
KEYWORDS JP 2002536973-A/10.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2342)
AUTHORS Valenzuela, D., Yuan, O., Hoffman, H., Hall, J., and Rapiejko, P.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patent: JP 2002536973-A 10 05-NOV-2002;
ALPHAGEN INC
COMMENT OS Homo sapiens (human)
PN JP 2002536973-A/10
PD 05-NOV-2002
PF 18-FEB-2000 JP 2000599860
PR 19-FEB-1999 US 60/120680, 23-APR-1999 US 09/298733 PR

17-AUG-1999 US 60/149639, 23-SEP-1999 US 60/155686 PR
01-OCT-1999 US 60/157247, 29-NOV-1999 US 60/167823 PR
29-NOV-1999 US 60/167822, 15-FEB-2000 US 60/182711 PI DARIO
VALENZUELA, OLIVE YUAN, HEIDI HOFFMAN, JEFF HALL, PETER PI RAPIEJKO
PC C12N15/09, A61K38/00, A61P3/10, A61P5/14, A61P11/00, A61P11/06, PC
A61P19/02,
PC A61P21/04, A61P25/14, A61P27/02, A61P29/00, A61P31/04, A61P31/10,
PC A61P31/12,
PC A61P31/18, A61P31/20, A61P31/22, A61P37/00, A61P37/06, C07K14/435,
PC C12N5/10,
PC C12P19/34, C12R1/91, C12N15/00, C12N5/00, A61K37/02
CC Secreted proteins and polynucleotides encoding them FH key
Location/Qualifiers
FT source 1. 2342
FT /organism='Homo sapiens (human)'.

FEATURES
source
1. 2342
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN

Query Match 33.4%; Score 692.2; DB 6; Length 2342;
Best Local Similarity 99.6%; Pred. No. 7.6e-91;
Matches 694; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 942 TATGAAGGCGCTGATTCGAAACCCAGCTAGCTGCTTCTCTCCCTGCGCCAGCTGCGGTGCT 1001
Db 1 TATGAAGGCGCTGATTCGAAACCCAGCTAGCTGCTTCTCTCCCTGCGCCAGCTGCGGTGCT 60
Qy 1002 CTCCTTCGCCACCGGGGCGCTGGGCATGTGGATCCCGCTTACCTGCAACCGCGCCAAAGT 1061
Db 61 CTCCTTCGCCACCGGGGCGCTGGGCATGTGGATCCCGCTTACCTGCAACCGCGCCAAAGT 120
Qy 1062 TGTGCAAGACAGCAGACAGCTGCAACAGCCGCGCTGCTGGGCGCAAGACAGCTCAT 1121
Db 121 TGTGCAAGACAGCAGACAGCTGCAACAGCCGCGCTGCTGGGCGCAAGACAGCTCAT 180
Qy 1122 CTTTGGGGCGCATACCTGCTTTACGGGATTTCTGGGCGTGTGCTGCGGGGCGAGGCCAC 1181
Db 181 CTTTGGGGCGCATACCTGCTTTACGGGATTTCTGGGCGTGTGCTGCGGGGCGAGGCCAC 240
Qy 1182 GCGCTGCTGCGCGCTGAAGACCCAGCGGGCGGACCCACTGCTGCTGCTGCTGCTGCTGCT 1241
Db 241 GCGCTGCTGCGCGCTGAAGACCCAGCGGGCGGACCCACTGCTGCTGCTGCTGCTGCTGCT 300
Qy 1242 GGGCTCTGCCATCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
Db 301 GGGCTCTGCCATCTTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 1302 CTATATCTGCTATCTTCTGCTGGGAGAGCTGCTGCTTCTAACTGGGCGCATCTGCGAGA 1361
Db 361 CTATATCTGCTATCTTCTGCTGGGAGAGCTGCTGCTTCTAACTGGGCGCATCTGCGAGA 420
Qy 1362 CATCTCTCATGTAGTGTATCTCCACGCGCGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCT 1421
Db 421 CATCTCTCATGTAGTGTATCTCCACGCGCGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCT 480
Qy 1422 CACTCTCCACTGCTGCTGGGAGCGCGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1481
Db 481 CACTCTCCACTGCTGCTGGGAGCGCGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 1482 GATCCGCGCAGACACTAAGGACTCCCGCTGCTGGGAGTCTCTGAGCGCTGGGCTAGCGCT 1541
Db 541 GATCCGCGCAGACACTAAGGACTCCCGCTGCTGGGAGTCTCTGAGCGCTGGGCTAGCGCT 600
Qy 1542 CATGCTCTGCGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1601
Db 601 CATGCTCTGCGCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 1602 CTTGCTCAGCGACCGCGCCAGGCTGAGCAGCCTG 1638
Db 661 CTTGCTCAGCGACCGCGCCAGGCTGAGCAGCAGGTG 697

RESULT 7
AX481748
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
CDS

AX481748
Sequence 6 from Patent WO2055701.
AX481748
AX481748.1 GI:22316582
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Curtis, R.A.
Human sugar transporter proteins, potassium channel proteins,
phospholipid transporter proteins and methods of use thereof
Patent: WO 0205701-A 6 18-JUL-2002;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..1587
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1587
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD44421.1"
/translation="MAGSDTAPFLSQADPDGVPVPGTGLPGSTGNPKSERPEVPDQ
EGLRITSPGRSLIIVAVLCVILNLYMDRFTVAGVLPDIEQFNIGDSSGLIQ
VFISYMWLAPVGLVGRNRYKVMCGIAFMSLVTLGSSFIPEHFWLLLTGLV
GVGEYSITPLIADLFVADQRMSLSIFPAIPVSGGLYIAGSKVMDVAGDHW
ALURVPGVAVLLFLVPRPPRGRVERHSDPLPLNLSWADRLAARNPFLS
SLGTFAPVTSALMAPAFLLRSRVILGETPCLPGDSCSSDLIFGLITGV
LGVGLVEISRLRSHNPADPLVATGLLSAPFLFLSLACARGSIIVATYIFIGE
TLLSNWATVADILLVWVTPRRSTAEAFQIVLSHLGLDAGSPYILGLISDRLRNP
PSFLSEFRALQSLMLCAFVGLGNAFLGTAIFIEADRRRAQLHVQGLHEAGSTDD
RIVVQGRSTRVPVNSVLI"

ORIGIN

Query Match 27.2%; Score 563.8; DB 6; Length 1587;
Best Local Similarity 64.3%; Pred. No. 3.2e-72;
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;
QY 282 GGGCCCCGGGCTCAGACGCCAAACGGCGAGCTTGGGCGCGGGGGGGGCGGCGGCGC 341
DB 126 GGACACGAGAGGGGCTGACGCGCATCACCGGCGCTCTCCCGCGCGTTCGAGT 185
QY 342 CGCATCTCAGCTTGGGCAACGTGCTCACTACCTGGACAGGTACACCGTGGCAGGCGT 401
DB 186 GCGGTGCTGTGTATCATCTCTCTGAACATGACGACCGCTTACCGTGGCTGGCGT 245
QY 402 CTTCTGGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGGCGCGGCGCTGTCAGTC 461
DB 246 CTTTCCGACATCGACAGCTTCTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAG 305
QY 462 AGTGTTCATCTGTAGCTTCATGTGGGTGCGCCCGTCTTGGCTACCTGGGCGACCGCTT 521
DB 306 CGTGTTCATCTCCAGTTACATGTGTGGCACCTGTGTGGTACCTGGGTGACAGGTA 365
QY 522 CAACAGGAGGTGATCTCAGCTGCGCGATTTTCTTGTGTGGCGGTACCTTCTCCAG 581
DB 366 CAATCGAGATATCTCATGTGCGGGGCAATGCTTCTGTGCTCCTGTGACATGGGTC 425
QY 582 CTCCTTCATTCCCAGCAGTACTTCTGGCTGTGGTCTGTCTCCCGGGGCTGGTGGCAT 641
DB 426 ATCTTCATCCCGGAGAGCATTTCTGGCTGTCTCTCTGACCGCGGCGCTGGTGGGT 485
QY 642 CGGGAGGCGAGCTACTCCACCATCGCCCGCATATCATTTGGGACCTCTTACCAGAA 701
DB 486 CGGGAGGCGAGTTATTCCACCATCGCGGCCACTCTCATTTGGCGGCACTCTTGTGCCGA 545

QY 702 CACGCTACGCTCATGCTGCTCGTCTTCTACTTGGCCATCCCACTGGGAGTGGCTGGG 761
DB 546 CCAGGGAGCCGGATGCTCAGCACTTCTACTTTGGCCATTCGGTGGGAGTGGTGGG 605
QY 762 CTACATTAATGCTGCTCAGCGTGAAGCAGGACCGGAGACTGGCACTGGGCATTTGGGGT 821
DB 606 CTACATTCAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGT 665
QY 822 GTCCCTCTCTGGGCGATGATCAGAGGAACACTCATCTCTCTCTCTCTCTCTCTCTCT 881
DB 666 GACACCGGCTTAGAGGTGGTGGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 725
QY 882 AAGGGGTCTATGCGCCAGCAGCTCGGGGA---CCAGCTCAAGGCCCGGAGACTCATATGCTCG 938
DB 726 AAGGGGCGCTGGAGCGCCACTCAGATTTGCGCACCCCTGAACCCACCCTCTCTCTCTCTCT 785
QY 939 AGATATGAGGCGCTGATTCGAAACCGGAGCTACGTCTCTCTCTCTCTCTCTCTCTCTCT 998
DB 786 AGATCTGAGGCGCTTGGCAAGAAATCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 845
QY 999 TCT 1058
DB 846 TGTGGCTCTTGTGACGGGCT 905
QY 1059 AGTTGTGCA---GAAGACAGCAGAGACGCTGGAACAGCCCGCTCTCTCTCTCTCTCTCTCT 1115
DB 906 CGTGTCTCTTGGGAGACCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 965
QY 1116 CCTCATCTTTGGGCGCATCCTGCTTTACGGGATTTCTGGGCGTGGTCTACGGGCGGAG 1175
DB 966 TCTCATCTTTGGACTCATCCTGCTCTGACCGGAGTCTCTGGGTGGGCTGGGTGGGA 1025
QY 1176 AGCCAGCGCTGCTGCTCGCTGAAAGACCCAGCGGCGGCGGACCCACTGCTGTGCTCGCTGG 1235
DB 1026 GATCAGCGCGCGCTCTCGGCACTCCACCCCGGCGTATCCCTCTCTCTCTCTCTCTCTCT 1085
QY 1236 CATGCTGGGCTCTGCGCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1295
DB 1086 CCTCTCTGGCTCTGCAACCT 1145
QY 1296 AGGAGCTTATCTGTATCTCTCTGGGAGAGCTGCTGTTTCTAACTGGGCGCATCAC 1355
DB 1146 GCGCACTTATATTTTCT 1205
QY 1356 TGCAGACATCTCTCATGTAGCTGTCATCCCAAGCGGCGGCGGCGGCGGCGGCGGCGGCG 1415
DB 1206 GCGCGACATCTCTGTAGCTGTCATCCCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1365
QY 1416 GAGCTTACCTCCCACTCTCTGGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1475
DB 1266 GATCGTCTGCT 1325
QY 1476 AGACCTGATCCGCGAGACACTAAGGACTCCCGCTCTGGGAGTCTCTGAGCTGGGCTA 1535
DB 1326 TGACCGCTCTGCGCGGAACTGGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1385
QY 1536 CCGCTCTCATGCTCTGCT 1595
DB 1386 CTCGCTCATGCTCTGCGGCTTGTGGGCACTGGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1445
QY 1596 GCTCTTCTTCTGTCAGCGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1642
DB 1446 CATCTTCTTCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1492

RESULT 8
BC006156
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

BC006156
Homo sapiens
partial cds.
BC006156
BC006156.1 GI:13544042

1836 bp
mRNA
linear
cDNA clone IMAGE:3627317,
spinster-like, mRNA

PRI 21-OCT-2003

| SOURCE | Homo sapiens (human) |
|----------|----------------------|
| ORGANISM | Homo sapiens |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

| AUTHORS | TITLE | JOURNAL MEDLINE PUBMED | REFERENCE | AUTHORS | TITLE | JOURNAL | REMARK COMMENT |
|---|---|---|--|--|---|---|-------------------|
| Klausner, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Krausner, R. D., Collins, P. S., Wagner, L., Shaenman, C. M., Schuler, G. D., Altschul, S. F., Zeeberg, B., Buetow, K. H., Schaefer, C. F., Bhat, N. K., Hopkins, K. F., Jordan, H., Moore, T., Max, S. I., Wang, J. J., Heieh, F., Diatchenko, L., Marasina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Bonaldo, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Usdin, T. B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullahy, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hulyk, S. W., Vitaloni, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Fahey, J. J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A. C., Shevchenko, Y., Bouffard, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schmutz, J., Myers, R. M., Butterfield, Y. S., Krzywinski, M. I., Skalska, U., Smalios, D. E., Schmerch, A., Schein, J. E., Jones, S. J. and Marra, M. A. | Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002) | 22389257 1247932 127 (bases 1 to 1836) | Strausberg, R. Direct Submission Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapsb@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) BC Cancer Agency, Vancouver, BC, Canada info@bcgsc.bc.ca | Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chiao, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhur, Parvaneh Saedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natsaja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra. | |

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 18 Row: h Column: 10.

FEATURES
source

```

Location/Qualifiers
1. .1836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="327317"
/tissue_type="Brain, neu"
/clone_libs="NIH MGC 19"
/lab_host="DH10B-R"
/notes="vector: pOT7"
<1. .1836
/gene="SPINL"
<1_xref="LocusID:83985"
<1. .1836
/gene="SPINL"
/codon_start=1
product="SPINL protein"

```

gene

CD5

```
<1. .15/5
/gene="SPINL"
/codon_start=1
/product="SPINL protein"
```

| | | | | |
|--------------|---------------------------|--|--------------------|--------------------|
| | | /protein_id="AAH06156.1" | | |
| | | /db_xref="GI:1354043" | | |
| | | /db_xref="LocusID:83985" | | |
| | | /translation="DTAPPLSQADDDPVPVGTGTLPGSTGNTGPKSEBPEVPDQEGLO RITGLSPGSAIIVAGLVLYNLYNMDRFTVAGVLPDLEQFNI GDSSELGLOTVPIS SNWLAIPVGYLGDYRNRYKLMCGGIAPWSLVLTGSSFFIPGEHFWMLLLTRGLVGVE ASYTGLVALLFLVADPORSNMLSIPIFYPAIPVSGLGYIAGSKVDKMDGMDHWWALRY TPGLGVAVLLFLVADPORSNMLSIPIFYPAIPVSGLGYIAGSKVDKMDGMDHWWALRY TAVAFYTGSLAWAPFLSVRLVGLGETPPCLPGDSCSSDSLI FGLITLFTGLVGLGVG LGVETSRRLHGNPAPFLVCAVLTGLGAPFLFLSLACARGSI VATYFIPIFETGLLS MKNWALVDLLVYVLPTRERSTAEAFQIVLSHLGDAQSPYLGLISDLRLRNWPSPL SFERALQFSIMLCAPVGLAGGAFLGTALIFIEADRRRAQLHVQGLLHEAGSTDDRLIV PQRGSTRVFPVASVLI" | | |
| | | 208..645 | | |
| misc_feature | | /gene="SPINL" | | |
| | | [note="AraJ; Region: Arabinose efflux permease [carbohydrate transport and metabolism]"] | | |
| | | /db_xref="CDD:COG2814" | | |
| ORIGIN | | | | |
| | Query Match | 27.2%; | Score 563.8; | DB 9; Length 1836; |
| | Best Local Similarity | 64.3%; | Pred. No. 3.1e-72; | |
| | Matches 879; Conservative | 0; | Mismatches 482; | Indels 6; Gaps 2; |
| QY | 282 | GGGCCCCCGCGCTCAGCAGCCAAACCGGCCAGCTTGGGCCCGGGCGGGGCGACGCCG | 341 | |
| DB | 114 | GGACGAGGAGGGCTGCAGCGCATCACCGCCGTGCTCCCGCCGTTCCGGCTCTCATAGT | 173 | |
| QY | 342 | CGCCATCCTCAGCTTGGGCAACGTGCTCAACTACCTTGGACAGGTACACCGTGGCAGCGT | 401 | |
| DB | 174 | GGCGGTGCTGTGCTACATCAATCTCTTGAACATACATGGACCGCTTTCACCGTGGCTGGCGT | 233 | |
| QY | 402 | CTTCTGGACATCCAGCAGCAGCTTTGGGGTCAAGGACCGAGGCGCGGCCCTCTGTCAGTTC | 461 | |
| DB | 234 | CTTCCCGACATCGAGCAGTCTTCAACATCGGGACAGTAGCTCTGGGCTCATCCAGAC | 293 | |
| QY | 462 | AGTGTTCATCTGTAGCTTCATGGTGGCTGCCCCCATCTTTCGGCTACCTGGGCGACCGTTC | 521 | |
| DB | 294 | CGTGTTCATCTCCAGTTACATGGTGTGGCACCTGTGTTGGCTACCTGGGTGACAGGTA | 353 | |
| QY | 522 | CAACAGGAGGTGATTCACAGCTGGCGCATTTTCTTCTGGTGGCGCGTCACTTCTCCAG | 581 | |
| DB | 354 | CAATCGGAAGTATCTCATGTGCGGGGGCAATGCGCTTCTGGTCCCTGGTGAACCTGGGTC | 413 | |
| QY | 582 | CTCCTTCATTTCCCGACGATCTCTGGCTGTGGTCTGTGTCCTCCGGGGCTGGTGGGCAT | 641 | |
| DB | 414 | ATCCTTCATCCCGGAGAGCATTTCTGGCTGTCTCTCTGTGACCCGGGGCTGGTGGGGT | 473 | |
| QY | 642 | CGGGAGGCGCAGCTACTCCACATCGCCCCCATATCATTTGGCGACCTCTTCCACCAAGAA | 701 | |
| DB | 474 | CGGGAGGCGCAGTTATCCACCATCGGCCCATCTCATTTGCCGACCTCTTTGTGCCGA | 533 | |
| QY | 702 | CAGGGTAGCTCATGTGTCGGTCTTCTATTCTGGCCATCCCATCTGGGGAGTGGCGCTGGG | 761 | |
| DB | 534 | CCAGCGGAGCGCGGATGCTCAGCATCTTCTACTTTGCCATTCGGGTGGGCGAGTGGTCTGGG | 593 | |
| QY | 762 | CTACATTACTGTGCTCCAGCGTGAAGCAGCAGCGGAGACTGGCACTGGGCACTGGCGGT | 821 | |
| DB | 594 | CTACATTGAGGCTCTAAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGT | 653 | |
| QY | 822 | GTCCCTGTCTGGGCATGATCACAGGAACHATCATCTCTCATTTCTGGTCCCGACCACTAA | 881 | |
| DB | 654 | GACACGGGTCTAGGAGTGGTGGCGTTCGTGCTGTCTCTGTGGTAGTGGCGAGCCGCC | 713 | |
| QY | 882 | AAGGGGTGATGCCGACAGCTCGGGGA---CCAGCTCAAGGCCCGGACCTCATTTGGCTCG | 938 | |
| DB | 714 | AAGGGGAGCGGTGGAGCGCCACTCAGATTTTGGCACCCCTGAACCCACCTCTGTGTGGGC | 773 | |
| QY | 939 | AGATATGAAGCCCTGATTCCGAACCGCAGCTACGCTCTTCTCTCCCTGGGCCACGTCGGC | 998 | |
| DB | 774 | AGATCTGAGGGCTCTGGCAAGAAATCTCTAGTTCGTCTGCTTCTCCCTGGGCTTCACTGC | 833 | |
| QY | 999 | TGTCCTCTTCGCCACGGGGGGCCCTGGGCATGTGGATCCCGCTCTACCTGCACCGCGGCCA | 1058 | |

| | | | | |
|-----------------------|-------|-----------------|----------------|-------------|
| Query Match | 27.2% | Score 563.8 | DB 9 | Length 2017 |
| Best Local Similarity | 64.3% | Pred. No. 3e-72 | | |
| Matches | 879 | Conservative 0 | Mismatches 482 | Indels 6 |
| Gaps | 2 | | | |

| | | | |
|----|------|--|------|
| QY | 282 | GGGCCCCGGCGCTCAGCAGCCCAACCGGCCAGCTTGGCGCGCGGGCGGGGCGAGCCGC | 341 |
| DB | 280 | GGACGAGAGGGGTGACGGCATCACGGGCTGTCTCCCGGCCGTTCCGGTCTCATAGT | 339 |
| QY | 342 | CGCCATCCTCAGCTTGGGCAACGTGCTCAACTACCTGGACAGGTACACCGTGGCAGCGGT | 401 |
| DB | 340 | GGCGGTCTGTGCTACATCAATCTCCTGAACTACATGACGCGCTTACCGTGGCTGGCGT | 399 |
| QY | 402 | CCTTCTGGACATCCAGCAGACACTTTTGGGTCAAGGACCGAGGCGCGGCTGTGTCAGTC | 461 |
| DB | 400 | CCTTCCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAGAC | 459 |
| QY | 462 | AGTGTTCATCTGTAGCTTCAATGGTGGCTGCCCCCATCTTGGGCTACCTGGGCGACCGGTT | 521 |
| DB | 460 | CGTGTTCATCTCCAGTTACATGGTGTGGACACCTGTGTGGCTACCTGGGTGACAGGTA | 519 |
| QY | 522 | CAACAGGAAGGTGATTTCTCAGCTGCGGCATTTTCTTCTGTGTGGCGCGTCACTTCTCCAG | 581 |
| DB | 520 | CAATCGGAAGTATCTCATGTGGGGGCAATTGCCCTTCTGTGTCCCTGGTGACACTGGGGTC | 579 |
| QY | 582 | CTCGTTCATTTCCAGCAGTAGTACTTCTGGCTGTCTGGTCTGTGTCGCGGGGCTGTGGGCAT | 641 |
| DB | 580 | ATCCTTTCATCCCGGAGAGCAATTTTGGCTGTCTCTCTGACCCGCGGCGCTGTGGGGGT | 639 |
| QY | 642 | CGGGAGGCGCAGCTPACTCCACATCGCCCCCACTATCATTTGGCGACACTTTCACCAAGAA | 701 |
| DB | 640 | CGGGAGGCGCAGTTATTCCACCATCGGCCCACTCTCATTTGCCACCTCTTTGTGGCGGA | 699 |
| QY | 702 | CACGCGTACGCTCATGCTGTCCGPTCTTACTTCGCCATCCCACTGGGCAAGTGGCCTGGG | 761 |
| DB | 700 | CCAGCGAGCGGATGCTCAGCACTCTTCTACTTTGGCCATTCGGTGGCGAGTGTCTGGG | 759 |
| QY | 762 | CTACATTACTGGCTCAGCGTGAAGCAGGACGCCGGAGACTGGGCACATGGGGATTGGCGGT | 821 |
| DB | 760 | CTACATTCGAGGCTCCAAGTGAAGGATATGGCTGAGAGCTGGGCACATGGGCTCTGAGGGT | 819 |
| QY | 822 | GTCCCTGTCTGGGCATGATCACAGAACACTCATCTCATTTCTGTGCCAGCCACTAA | 881 |
| DB | 820 | GACACGGGTCTAGGAGTGTGGCGGTCTCTGCTGCTTCTGTGTAGTGGGAGCGCCGC | 879 |
| QY | 882 | AAGGGTCTATGCCAGCAGCTCGGGGA---CAGCTCAAGGCCCGGACCTCATGGCTCCG | 938 |
| DB | 880 | AAGGGAGCGGTGGAGCGCCACTCAGATTTGGCCACCCCTGAACCCCACTCGTGTGGGC | 939 |
| QY | 939 | AGATATGAAGGCCCTGATTCGAACCGCAGCTAGTCTTCTCTCCCTGGCCAGCTGGC | 998 |
| DB | 940 | AGATCTGAGGGCTCTGGCAAGAAATCCTAGTCTGTGCTCTGTCTTCCCTGGGCTTCACTGC | 999 |
| QY | 999 | TGTCTCCTTCGCCACGGGGGCCCTGGGCATGTGGATCCCGCTCTACTCTGACACCGCGCCA | 1058 |
| DB | 1000 | TGTGGCCTTTGTACGGGCTCCCTTGGCTCTGTGGGCTCTGGGCATTCCTGTGCTGCTCCG | 1059 |
| QY | 1059 | AGTTGTGCA---GAAAGACAGCAGAGAGCTGCAACAGCCCGCCCTGTGGGCCCAAGGACAG | 1115 |
| DB | 1060 | CGTGGTCTTTGGGAGAGCCCACTCCCTGCCCTTCCGGAGATCTCTGTCTTCTCTGTGACAG | 1119 |
| QY | 1116 | CCTCATCTTTGGGGCCATCATCTGCTTTACGGGATTTCTGGGCGTGTGTACGGGGGACAG | 1175 |
| DB | 1120 | TCTCATCTTTTGAATCATACCTGCCCTGACCGGAGTCTCTGGGTGGGCGCTGGGTGTGGA | 1179 |
| QY | 1176 | AGCCACCGCTGTGTCGGCTCTGAAGACCCAGCGGGCGGACCCACTGTGTGTGTCGCTGGG | 1235 |
| DB | 1180 | GATCAGCGCGGCTCGGCCACTTCCAAACCCCCGGGCTGATCCCTGTGTGTGTGCCACTGG | 1239 |
| QY | 1236 | CATGCTGGGCTCTGCCATCTTTCATCTGCCCTGATCTTCTGTGGCTGCCAAGACGACATCGT | 1295 |

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mc@nhgri.nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maekari, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
 Series: IRAC Plate: 140 Row: m Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Genomescan gene
 prediction, Similarity but not identity to protein.

FEATURES

source

1. .2049
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:70841 IMAGE:6161778"
 /tissue_type="Skin, melanotic melanoma."
 /clone_lib="NIH MGC_72"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

gene

1. .2049
 /gene="SPINL"
 /db_xref="LocusID:83985"
 123. .1709
 /gene="SPINL"
 /codon_start=1
 /product="SPINL protein"
 /protein_id="AAH65235.1"
 /db_xref="GI:40807118"
 /db_xref="LocusID:83985"
 /translation="MAGSDTAPFLSQADDPDGPVPGTGLPGSTGNPKSERPEVPDQ
 EGLORTGLSPGRSALIVAVLCVILNLYMDRTVAGLPDIEOFENIGDSSGLIOT
 VFTSSVMVLAPVGYLDGRNRYLMCGGIAFWSLVTGLSSPIRGEHFWLLLTGLV
 GVGEASYSIATPLIADLVADQSRMSLIFYPATPVGSGLYIAGSKYKVDNAGDWHV
 ALRVTGLGVFVLLFLVVRPPRGVSRVHLNPTNSWADRLARNPSFLS
 SLGFTAVAFVTSGLAWAPALLRSRVILGETPPCLPGDSCSSSDSLIFGLITCLTV
 LGVGLGEISRLRHSNRPADPLVCATGILGAPFLFLACARGSI VATYIFIFIGE
 TLLSMWAI VADILLYVVIPTRRSTAEAFQIVLSHLGDAGSPYILGLISDLRRNP
 PSFLSFRALQFSLMLCAFVGLGGAFLGTALFIEADRRRAQLHVQGLLHEAGSTDD
 RIVVPGRSTRVPVAVSLI"

CDS

393. .776
 /gene="SPINL"
 /note="Sugar tr; Region: Sugar (and other) transporter"
 /db_xref="CDD:pfam00083"

misc_feature

ORIGIN

Query Match 27.2%; Score 563.8; DB 9; Length 2049;
 Best Local Similarity 64.3%; Pred. No. 3e-72;
 Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;
 QY 282 GGGCCCCGGCTCAGCAGCCCAACCGGCAGCTTGGCCGCGGGGGGGGCGAGCCGC 341
 DB 248 GGACAGAGAGGGCTGCAGGGATCACCGGCTGTCTCCGCGCGTTCGGCTCTCATGT 307
 QY 342 CGCCATCTCAGCTTGGGCAACGCTGCTCAACTACTTGGACAGTACACCGTGGCAGGCGT 401
 DB 308 GCGGTGCTGTGTACATCAATCTCTGAATACATGAGACCGCTTCAACGCTGGCGT 367
 QY 402 CTTTCTGGACATCCAGCAGCTTTGGGCTCAAGNACGAGGGCGGCGCTCTCTCAGTC 461
 DB 368 CTTTCCGACATCGAGAGTTCTTTCAACATCGGGGACAGTAGTCTCTGGGCTCATCGAC 427

QY 462 AGTGTTCATCTCTAGCTTCATGGTGGCTGCCGCCCATCTTTCGGCTACCTGGGGGACGCGTT 521
 DB 428 CGTGTTCATCTCCAGTTACATGGTGTGGCACCTGTGTGGTACCTACCTGGGTGACAGGTA 487
 QY 522 CAACAGGAGGTGATTCTCAGCTGGGGCATTTTCTTCTGGTGGCGCGTCACCTTCTCCAG 581
 DB 488 CAATCGGAGTATCTCATGTGGGGGGCATTCGCTTCTGGTCCCTGGTGGACACTGGGGTC 547
 QY 582 CTCCTTCATTTCCCAAGCAGTACTTCTGGCTGTGTGGTCTCTCCCGGGGCTGGTGGGCA 641
 DB 548 ATCTCTTCATCCCGGAGAGCAATTTCTGGCTGTCTCTCTGACCCCGGGGCTGTGGGGGT 607
 QY 642 CGGGAGGCGCAGTACTACTCACCATCGCCCCCACTATCATATTGGCGACCTCTTTCACCAAGAA 701
 DB 608 CGGGAGGCGCAGTATTATCCACCATCGGCCCATCTCTCATTTGCGGACCTCTTTTGGCCGA 667
 QY 702 CACGGTATAGCTCATGTCTCGCTCTTCTACTTTCGCCATCCCATCTGGGAGTGGCTGG 761
 DB 668 CCAGCGGAGCGGATGCTCAGCACTTCTACTTTGCCATTCGGTGGGAGTGGTCTGG 727
 QY 762 CTACATTACTGGCTCCAGCGTGAAGCAGGCGAGACTGGSCACTGGSCATTTGGGGT 821
 DB 728 CTACATTGCAAGCTCCAAAGTGAAGATATGCTGAGACTGGCACTGGGCTCTGAGGCT 787
 QY 822 GTCCCCCTGTCTGGGCATGATCAGAGGAACACTCATCTCTCATTTCTGGTCCCAAGCCACTAA 881
 DB 788 GACACCGGGTCTAGGAGTGGTGGCGTTCTGCTGTCTTCTTGGTAGTGGCGGAGCCGCC 847
 QY 882 AAGGGTCTATGCGCGACCACTCGGGGGA--CCAGCTCAAGGCCCGGAGCTCATGTGCTCG 938
 DB 848 AAGGGAGCGGTGGAGCGCCACTCAGATTGGCACCCCTGAACCCCACTCTGTGTGGGC 907
 QY 939 AGATATGAAGCCCTGATTTCGAAGCCGAGTACGTCTTCTCTCTCTGGGCACTCTGGC 998
 DB 908 AGATCTAGGGGCTCTGGCAAGAAATCTAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 967
 QY 999 TGTCTCTCTTGGCCAGCGGGGCTGGGCATGTGGATCCGGCTCTACCTTGCACCGGCCCA 1058
 DB 968 TGTGGCTCTTGTACGGGCTCTCTGGGCTCTCTGGGCTCTCTGGGCTCTCTCTCTCTCT 1027
 QY 1059 AGTTGTGCA---GAAGACAGCAGAGAGAGCTGCAACAGCCCGGCTGTGGGGCCAGGACAG 1115
 DB 1028 CGTGTCTCTGGGAGACCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1087
 QY 1116 CTTCTCTCTTGGGGCCATCACTGTCTTTCAGGGAATTTCTGGGCGTGTGTCAAGGGGGCAG 1175
 DB 1088 TCTCATCTTTGGACTCATCACTGTCTGACCGGAGTCTCTGGGTGTGGGCTCTGGGTGTGGA 1147
 QY 1176 AGCCACGCGTGTGGCGGCTGAAGACCCAGCGGGCGGACCCACTGTGTGTGGCGTGG 1235
 DB 1148 GATACGCCCGGCTCTCCGCACTCAACCCCGGGCTGATCCCTGTGTCTGTGCCACTGG 1207
 QY 1236 CATGCTGGGCTCTGCACTCTTTCATCTGCTGTGATCTTCTGTGGCTGTGCAAGAGCAGCATCG 1295
 DB 1208 CCTCTGGGCTCTGCACT 1267
 QY 1296 AGGAGCTATATCTGTATCTTCTGTGGGAGACCGTGTCTGTCTTCTAACTGGGCCATCAC 1355
 DB 1268 GGCCACTTATATTTTCATCTTTCATTTGGAGAGACCTCTCTCTCATGAACCTGGGCCATCG 1327
 QY 1356 TGCAGACATCTCATGTACGTGGTCACTCCCGCGGGCGGCCACTGTCCGTGGCTTGTCA 1415
 DB 1328 GGCGACATCTGT 1387
 QY 1416 GAGCTTCACTCCCACTGTGTGGGGAGCGCGGGAGCCCTTCACTCAATTGGCTTTATCTC 1475
 DB 1388 GATCGT 1447
 QY 1476 AGACTGATCGGCGAGCAGTAAAGGACTCCCGCTCTGGGAGTCTCTGAGCTTGGGGTA 1535
 DB 1448 TGACCGCTCGCGGGAAGTGGGCCCTCTCTTCTTGTCCGAGTTCGGGCTCTGCACTT 1507

Qy 1536 CGCGCTCATGCTCTGCCCTTTCGTCTGCTGCTGCGGCGGCGATGTTCTTCTCGCCACTGC 1595
Db 1508 CTCGCTCATGCTCTGCCGCTTGTGTGGGCGCACTGGCGGCGCGACGCTTCTCGGGCAGCGC 1567
Qy 1596 GCTCTTCTTCGCTCAGCGACCGCGCCAGGCGCTGAGCAGCACCTGGGG 1642
Db 1568 CATCTTCATGAGCGCGACCGCGCGGCGGCGACGCTGCACGTGCAGG 1614

RESULT 11
BC008325 2050 bp mRNA linear PRI 29-JUN-2004
LOCUS Homo sapiens spinster-like, mRNA (cDNA clone MGC:15767
DEFINITION IMAGE:3501826), complete cds.
ACCESSION BC008325
VERSION
KEYWORDS
SOURCE MGC.
ORGANISM Homo sapiens (human)

REFERENCE
AUTHORS Strausberg, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 2050)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NTH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Joeh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skalska, Duane Smalls, Jeff Scott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 23 Row: b Column: 17.
Location/Qualifiers

FEATURES

source 1. .2050
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:15767 IMAGE:3501826"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH_MGC_19"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
1. .2050
/gene="SPINL"
/db_xref="LocusID:83985"
203. .1789
/gene="SPINL"
/codon_start=1
/product="spinster"
/protein_id="AAH08325.1"
/db_xref="GI:14249892"
/db_xref="LocusID:83985"
/translation="MAGSDTAPFLSQADDDPDPVPGTPTGPGTGNPKSEPEVPDQ
EGQRITGLSPGRSALIVVLCYINILNMYDRFTVAGVLPDIEOFFNIGDSSGLIQ
VFISYVVLAPVFGYLDVNRKYLMCGGIAPFWSLVTLGSSFI PGEHFWLLLLTRGLV
GVGEASSTIAPTLLIADLFVADORSMLSI FYFAIPVSGSLGYIAGSKVKDMAGDWHH
ALRVTPGLVAVVLLFLVREPRGAVRHSRVLGETPECPDSCSSDSLIFFGLITCTGV
SLGFTVAFVTVGSLAWAPFLRSRVLTGETPECPDSCSSDSLIFFGLITCTGV
LGVLGVEI SRRLHNSPRADPLVCATGLGSAFLFLSLACARGSIVATYIFFIGE
TLSSMMWAI VADILLVYVIVPTRTAEAFQIVLSHLGADGASPYLIGLISRLRNWP
PSFLSEFRALQFSLMLCAFYGALGGAFLGTAIFIEADRRRAQLHVQGLLHAGSTDD
RIVVPQGRSTRVPVASVLI"
ORIGIN
Query Match 27.2%; Score 563.8; DB 9; Length 2050;
Best Local Similarity 64.3%; Pred. No. 3e-72;
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;
Qy 282 GGGCCCCGGCGCTCAGCAGGCCAAACCGGCCAGCTTGGGCGCGGGGGGCGCGCGC 341
Db 328 GGACCAGAGGGGCTGCAGCGCATCACCGCGCTCTCTCCGCGCGCTTCGGCTCTCATAGT 387
Qy 342 CGCCATCTCAGCTTGGGCGACGTCCTCACTACCTGACAGGTACACCGTGGCAGCGGT 401
Db 388 GGGCGTCTGTGTACATCAATCTCTGAACTACATGGACCGCTTACCGGCTGGCGGT 447
Qy 402 CTTTCTGGACATCCAGCAGCACCTTTGGGGTCAAGGACCGAGGCGCGCGCTTCTCCAGTC 461
Db 448 CTTTCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAGAC 507
Qy 462 AGTGTTTCATCTGAGCTTTCATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 521
Db 508 CGTGTTCATCTCCAGTTACATGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 567
Qy 522 CAACAGGAGGTGATCTTCAGCTGGCGCATTTTCTGCTGGTGGCGCTTCTCCAGTC 581
Db 568 CAATCGGAAGTATCTCATGTGGGGGCGCATTTGCTTCTGCTGGTGGACACTGGGGGTC 627
Qy 582 CTCCTTCATTCGCCAGCAGTACTTCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 641
Db 628 ATCTTTCATCCCGGAGAGCATTTCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 687
Qy 642 CGGGGAGGCGAGTACTTCCACCATCGCCCCCATATCATTTGGGACCTCTTTCACCAAGAA 701
Db 688 CGGGGAGGCGAGTATTTCCACCATCGCGGCCACTCTCATTTGGGACCTCTTTTGGGCGCA 747
Qy 702 CACGGTTCAGCTCATCTGCTGCTCTCTACTTGGCCATCCACTGGGCGAGTGGCGCTGGG 761
Db 748 CCAGCGAGCGCGATGCTCAGCATCTTCTACTTTGGCATTTCCGGTGGGCGAGTGGTCTGGG 807
Qy 762 CTACATTACTTGGCTCCAGCGCTGAAGCAGGCGGAGACTGGCACATGGGCACTTGGCGGT 821
Db 808 CTACATTGAGGCTCCAAAGTGAGGATATGGGTGGAGATGGGCACTGGGCTCTGAGGGT 867
Qy 822 GTCCCTGTCTCGGCGATGATCA CAGGAACACTCATCTCTCATTTCTGGTCCCGACCACTAA 881

| | | | |
|----|------|---|------|
| Db | 990 | GACACCGGGTCTAGGAGTGTGGCCGTTCTGCTGCTTTCTTGGTAGTGGCGGAGCGCCG | 1049 |
| Qy | 882 | AAGGGGTATGCCACACGAGTCTGGGGG- --CCAGCTCAAGCGCCGGACCTCATGGGTCGG | 938 |
| Db | 1050 | AAGGGAGCGTGGAGCGCCACTCAGATTGGCCACCCTGAAACCCACACTCGTGTGGG | 1109 |
| Qy | 939 | AGATATGAAGGCCCTGATTCGAAACCGAGTACGTCTTCTCTCTCTGACCCACGTCGCG | 998 |
| Db | 1110 | AGATCTGAGGGCTCTGGCAAGAAATCTAGTTTTCGTCTCTCTCTCTCTCTCTCTCTCT | 1169 |
| Qy | 999 | TGTCCTCTTCGCGACCGGGGCGCTGGGCAATGTGGATCCCGCTCTACCTGTCACCGCGCCA | 1058 |
| Db | 1170 | TGTGGCCTTTGTACGGGCTCCCTGGCTCTGTGGGCTCCGGCAATCTCTGTGCTGTTCCG | 1229 |
| Qy | 1059 | AGTTGTGCA- --GAAAGACACGAGACGTGCAACAGCGCCGCTGTGGGCGCAAGACAG | 1115 |
| Db | 1230 | CGTGGTCTTGGGAGACCCACCGCTGCCCTCCGGAGACTCTCTCTCTCTCTCTCTCTCT | 1289 |
| Qy | 1116 | CCTCATCTTTGGGGCCATCACTGCTTTACGGGATTTCTGGGGGTGTCTACGGGGGCGAG | 1175 |
| Db | 1290 | TCTCATCTTTGGACTCATCACTGCCTGACCGGAGTCTCTGGGTGGGCTGGGTGGA | 1349 |
| Qy | 1176 | AGCCACGCGTGTGCGCCTGAAGAACCCAGCGGCGCGAACCACTGGTGTGCGCGTGGG | 1235 |
| Db | 1350 | GATCAGCGCGCGCTCGGCCACTCAACCCCGGGCTGATCCCTGGTCTGTGCCACTGG | 1409 |
| Qy | 1236 | CATGCTGGGCTCTGCCATCTCATCTGCCCTGATTTCTGGTGTGCCAAGACGACATCGT | 1295 |
| Db | 1410 | CTCTCTGGGCTCTGCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 1469 |
| Qy | 1296 | AGGAGCCTATCTGTATCTTCTGTGGGGAGACGCTGCTCTTTTCTTAACGTGGGCCATCAC | 1355 |
| Db | 1470 | GGCCACTTATTTTCACTTCAITGGAGAGACCTCTCTCTCATGAACTTGGGCCATCGT | 1529 |
| Qy | 1356 | TGCAGACATCCTCATGTAGTGTATCTCCACGCGCGCGCCACTGCCGTGGCCTTGCA | 1415 |
| Db | 1530 | GGCCGACATTCTGCTGTATCGTGTGTATCCCTACCCGACGCTCCACCGCGAGGCTTCCA | 1589 |
| Qy | 1416 | GAGCTTCACTCCCACTCTGTGGGGAGCGCGGAGGCCCTACTCATTTGGCTTTATCTC | 1475 |
| Db | 1590 | GATCGTGCTGCCACCTCTGTGGGTGATGTGGGAGCCCCCTACTCATTTGGCCTGATCTC | 1649 |
| Qy | 1476 | AGACCTGATCCGCCAGACACTAAGGACTCCCCGCTCTGGGAGTCTCTGAGCCTGGGCTA | 1535 |
| Db | 1650 | TGACCGGCTGCGCGGAACTTGGCCCCCTCTTCTTGTTCGAGTTCCGGGCTCTGAGTT | 1709 |
| Qy | 1536 | CGGCTCATGCTCTGCCCTTTCGTGTGTCTCTGGGCGGATGTTCTTCTCTGCCACTGC | 1595 |
| Db | 1710 | CTCGCTCATGCTCTGGCGTTTGTGGGACACTGGGCGGGCGAGCCTTCTCTGGGACCGC | 1769 |
| Qy | 1596 | GCTCTTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACTCGGGG | 1642 |
| Db | 1770 | CATCTTCATTGAGCGCGACCGCGCGGGGACAGCTGCACTGCAGTGCAG | 1816 |

[illegible]

Db 1075 TCTGAGGCTCTGGCAAGAAATCTAGTTCTGCTGCTTCTCCCTGGGCTTCACTGCTGT 1134
Qy 1002 CTCCTTGGCCACGGGGCCCTGGGCACTGTGATCCCGCTCTACCTGACACCGCGCCCAAGT 1061
Db 1135 GGCCTTTGTACGGGCTCCCTGGCTCTGTGGGCTCCGGCAATCTCTGCTGGTTCGGCGT 1194
Qy 1062 TGTGCA---GAAGACAGAGAGAGCTGCAACAGACCGCCCTGTGGGGCCAAAGACAGCT 1118
Db 1195 GGTCTTTGGGAGAGACCCACCTGCCCTTCCCGAGAGACTCTGCTCTCTCTGACAGTCT 1254
Qy 1119 CATCTTTGGGGCCATCACTCTTTACGGGATTTCTGGGATTTCTGGGCTGTACAGGGGAGAGC 1178
Db 1255 CATCTTTGGACTCATCACTGCTTGACCGGAGTCTGGGTGGGCTGTGGGTGGAGAT 1314
Qy 1179 CACCGCTGTGGCCCTGAAAGACCCAGCGGGCCGACCACTGTGTGTGTCGGTGGGCAT 1238
Db 1315 CAGCGCGGCTCCGCCACTCCACACCCCGGGCTGATCCCTGTCTGTGCTGCTGCTGCT 1374
Qy 1239 GCTGGGCTCTGCCATCTTCATCTGCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
Db 1375 CTTGGGCTCTGCCACTTCT 1434
Qy 1299 AGCTATATCTGATCTCTGCTGGGAGAGCGCTGCTGCTTCTTCTTCTTCTTCTTCT 1358
Db 1435 CACTTATATTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1494
Qy 1359 AGACATCTCTATGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1418
Db 1495 CGACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1554
Qy 1419 CTTTACCTCTCCACTCTGCTGGGGAGCGCGGGAGCCCTTCTTCTTCTTCTTCTTCT 1478
Db 1555 CGTGTCTGCCACTCTGCTGGGTGATGCTGGGAGCCCTTCTTCTTCTTCTTCTTCT 1614
Qy 1479 CTTGATCCGCGAGAGCACTAAGACTCTCCGCTCTGGGAGTCTGAGGCTGGGCTAGCG 1538
Db 1615 CCGCTCTGCGCGGAACCTGGCCCTCTCTTCTTCTGCGAGTCTGGGCTCTGCACTTCT 1674
Qy 1539 GCTCATGCTCTGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1598
Db 1675 GCTCATGCTCTGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1733
Qy 1599 CTTCTTCTGTCAGCGACCGCGCCAGGCTGAGCAGCAGCTGCGGG 1642
Db 1734 CTTTATTGAGCCGACCGCGCGGCGACAGCTGACGTGACAG 1777

RESULT 15

AX924785
LOCUS 2201 bp DNA linear PAT 19-DEC-2003
DEFINITION Sequence 240 from Patent EP1352962.
ACCESSION AX924785
VERSION AX924785.1 GI:40242919

KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified.

REFERENCE

1. Olsen, H.S., Yu, G.L., Endress, G.A., Bednariak, D.P., Carter, K.C.,
Feng, P., Soppet, D.R., Young, P.E., Duan, D.R., Florence, K.A.,
Greene, J.M., Fischer, C.L., Hu, J.S., Ruben, S.M., Ebner, R.,
Brewer, L.A., Ferrie, A.M., Moore, P.A., Ni, J., Shi, Y., Lafleur, D.W.,
Li, Y., Zeng, Z., Kyaw, H. and Rosen, C.A.
186 human secreted proteins

TITLE

JOURNAL Patent: EP 1352962-A 240 15-OCT-2003;

FEATURES

Location/Qualifiers
1. .2201
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 26.2%; Score 542.8; DB 6; Length 2201;
Best Local Similarity 64.3%; Pred. No. 3.3e-69;
Matches 877; Conservative 1; Mismatches 478; Indels 8; Gaps 4;
Qy 285 CCCCGGCGCTCAGCAGCCCAACCGGCGAGTGGGCGCGGGGGGGGAGCGCGCGC 344
Db 416 CCAGGARGGGCTGACGGGCATCAGCGGCTGTCTCCCGCGGCTTCCGGCTCTCATAGTGGC 475
Qy 345 CATCTCAGCTTGGGCAACGTGTCAACTACCTGACAGGTACACCGTGGCAGGCTCCT 404
Db 476 GGTGCTGTGTACATCAATCTCTGAACATACAGCAGCTTCAACGTTGGTGGCTGCT 535
Qy 405 TCTGGACATCCAGCAGCACTTTGGGTCAGAGGACCGAGGCGCCGCTCTGCTGCACTCAGT 464
Db 536 TCCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCAGACGCT 595
Qy 465 GTTCATCTGTAGCTTCAATGTGGTGGCTGCCCATCTTGGCTACCTGGGCGAGCCGCTTCAA 524
Db 596 GTTCATCTCCAGTTACATGGTGTGGCACTGTGTGGCTTACCTGGGTGACAGGTACAA 655
Qy 525 CAGGAAGGTGATTTCTCAGCTGCGGCATTTTCTGTGTCGCGCTCACCTTCTCCAGTTC 584
Db 656 TCGGAAGTATCTCATGTGCGGGGCACTTGTCTTGTCTGCTGCTGCTGCTGCTGCTGCT 715
Qy 585 CTTTCTTCCCGCAGCAGTACTTCTGGCTGTGTGCTCTTCCCGGGGGCTGGTGGGCTCGG 644
Db 716 CTTTCTTCCCGGAGAGCATTTCTGGCTGTCTCTCTGACCCCGGGGCTGGTGGGCTCGG 775
Qy 645 GAGGCGCAGCTACTCAGCATTCGCCCTCCACTATCATTTGGGAGCCTCTTACCAAGAACAC 704
Db 776 GAGGCGCAGTTATTTACCATCGCGCCACTCTCATTTGCGGACCTCTTTGTGGCGGACCA 835
Qy 705 GCGTAGCTCATGCTGCTGCTCTTCTTCTGCGCATCCACTGGGAGTGGCTGGGCTGTA 764
Db 836 GCGGAC--CGGATGCTCAGCATCTTCTTCTTCTGCGCATTCGCGTGGGAGTGGCTGGGCTA 894
Qy 765 CATTTACTGGCTCCAGCTGAGGAGCGGAGAGCTGGGAGCTGGGAGCTTGGGCTGGCTGTC 824
Db 895 CATTTGAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGGAGCTGGGCTCTGAGGCTGAC 954
Qy 825 CCTGTCTGGGAGATGATCAAGAACTCATCTCTCATTTGCTGCTCCAGGCACTAAAG 884
Db 955 ACCGGGCTAGGAGTGGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
Qy 885 GGGTCATGCGCAGCAGCTCGGGA---CCAGCTCAAGGCGCGGAGCTCATGGCTCCGAGA 941
Db 1015 GGGAGCGGTGAGCGGCGACCTCAGATTTGCGACCCCTGAAACCCACCTCTGTTGGTGGG 1074
Qy 942 TATGAAGGCGCTGATTTGAAACCGCAGCTACGCTTCTTCTCTCTGCGGAGCTGCGGCTGT 1001
Db 1075 TCTGAGGCTCTGGCAAGAAATCTTAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
Qy 1002 CTCCTTGGCCACGGGGCCCTGGGAGTGGATCCCGCTCTACTGTCAGTGGCTTCACTGCTGT 1061
Db 1135 GGCCTTTGTACCGGCTCCCTGGCTCTGTGGGCTCCGGCATTCCTGCTGCTGCTGCTGCTGCT 1194
Qy 1062 TGTGCA--GAGACAGCAGAGAGCTGCAACAGCCCGCTCTGTTGGGCGGAGGAGCAGCT 1118
Db 1195 GGTCTTGGGAGAGACCCCTGCTTCCCGGAGACTCTGCTCTTCTTCTTCTTCTGACAGTCT 1254
Qy 1119 CATCTTTGGGGCCATCACCTGCTTTTACGGGATTTCTGGGCGTGGTCAAGGGGCGAGGAGC 1178
Db 1255 CATCTTTGGACTCATCACTGCTGCTGACCGGAGTCTCTGGGTGGGCTGGGTGGTGGAGAT 1314
Qy 1179 CAGCGCTGTGCTCGCTGAAAGACCCAGCGGGCGGAGCCCACTGGTGTGCTGCTGGGCTAG 1238
Db 1315 CAGCGCGGCTCGGCACTTCCAAACCCCGGGCTGATCCCTGCTGCTGCTGCTGCTGCTGCT 1374
Qy 1239 GCTGGGCTCTGCCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
Db 1375 CTTGGGCTCTGCAACCCCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1434

| | | | |
|----|------|--|------|
| Qy | 1299 | AGCTATATCTGTATCTTCGTGGGAGAGCGTGTGTTTCTTAACCTGGGCCATCACTGC | 1358 |
| Db | 1435 | CACTTATATTTTATCTTATTGGAGAGACCCCTCTGTCCATGAACCTGGGCCATCGTGGC | 1494 |
| Qy | 1359 | AGACATCCTCATGTACGTGGTCAATCCACGCGGGCGGCACTGCCGTGGCCTTGCAGAG | 1418 |
| Db | 1495 | CGACATTTCTGTGTAGTGTGTATCCCTACCCGAGGCTCCACCGCGAGGCCCTTCCAGAT | 1554 |
| Qy | 1419 | CTTCACCTCCACCTGCTGGGGAGCGCCGGGAGCCCTACCTCATTTGCTTTATCTCAGA | 1478 |
| Db | 1555 | CGTGTGTCCTCCACCTGCTGGGTGATGCTGGGAGCCCTACCTCATTTGCTGTCTCTGA | 1614 |
| Qy | 1479 | CCTGATCCGCCAGAGCACTAAGGACTTCCCGGCTCTGGGAGTTCTTGAGCCCTGGGCTACGC | 1538 |
| Db | 1615 | CGCCCTGGCGCGAACTGGCCCCCTCTTCTTGTCCGAGTCCGGGCTCTGCAGTTCTC | 1674 |
| Qy | 1539 | GCTCATGCTCTGCCCTTTCTGTGTGTCTGGGGGCAATGTTCTTCTCGCCACTGGGCT | 1598 |
| Db | 1675 | GCTCATGCTCTGGCGTTTGTGGGGCACTGGGGGCGCACT-TTCCTGGGCACCGCCAT | 1733 |
| Qy | 1599 | CTTCTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACTGGGGG | 1642 |
| Db | 1734 | CTTCATTGAGGCCGACCGCGCGGGCAACAGCTGCAGTGCAGG | 1777 |

Search completed: October 23, 2004, 11:01:11
Job time : 8813 secs

PR 12-APR-2001; 2001US-0283443P.
PR 23-APR-2001; 2001US-0285754P.
PR 24-APR-2001; 2001US-0286096P.
PR 03-MAY-2001; 2001US-0288353P.
PR 17-MAY-2001; 2001US-0291703P.
PR 31-MAY-2001; 2001US-0294834P.
PR 20-JUN-2001; 2001US-0295695P.
PR 21-JUN-2001; 2001US-0298845P.
PR 05-JUL-2001; 2001US-0303242P.
PR 13-AUG-2001; 2001US-0311981P.
PR 16-AUG-2001; 2001US-0312858P.
PR 17-AUG-2001; 2001US-0313280P.
PR 29-AUG-2001; 2001US-0315614P.
PR 17-SEP-2001; 2001US-0322818P.
PR 25-FEB-2002; 2002US-00322818.
XX (CURA-) CURAGEN CORP.
XX
XX
PI Alsobrook JP, Anderson DW, Ballinger RA, Boldog FI, Burgess CE;
PI Casman SJ, Ellerman KE, Gangolli RA, Gerlach VI, Gilbert JA;
PI Gorman L, Guo X, Gusev VY, Kekuda R, Li L, Liu X, Malyankar UM;
PI Miller CE, Millet I, Padigaru M, Patturajan M, Pena CE, Peyman JA;
PI Rastelli L, Shenoy SG, Shimkets RA, Smithson G, Spytek KA, Stone DJ;
PI Taupier RU, Tchernev VT, Vernet CAM, Zerhusen BD;
XX
DR WPI; 2002-698672/75.
DR P-PSDB; ADH48764.
XX
XX
PT New NOVX polypeptides or polynucleotides, useful for preventing or
PT treating disorders or syndromes e.g., atherosclerosis, hypertension,
PT obesity or cancer.
XX
XX
PS Claim 8; Page 129-130; 923pp; English.
XX
XX
CC The present invention relates to novel human NOVX proteins, where X is
CC any number from 1 to 91 and their coding sequences. The proteins and
CC coding sequences are useful for preventing or treating disorders or
CC syndromes e.g. atherosclerosis, hypertension, obesity or cancer. NOV19 is
CC a spinster-like protein and its coding sequence maps to chromosome 17.
XX
SQ Sequence 2071 BP; 341 A; 681 C; 685 G; 364 T; 0 U; 0 Other;

Query Match 100.0%; Score 2071; DB 6; Length 2071;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2071; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCGCCCGATCCGGGCGGCATGATGCTGCTGGAATGCGCCCTCGGCGGGCG 60
DB 1 CCCCCCGCCCGATCCGGGCGGCATGATGCTGGAATGCGCCCTCGGCGGGCGG 60

QY 61 GCGGCGCGGAGAGGAGGCGGACGCGGAGCGCGCGCGCGCGCGGGGGCGGAG 120
DB 61 GCGGCGCGGAGGAGGAGGCGGACGCGGAGCGCGCGCGCGCGCGGGGGCGGAG 120

QY 121 CGAGGGGCTGGCGGTAGCGTTCCTGCGGGGCGCGGGGCGCGGGGCGGCTGCG 180
DB 121 CGAGGGGCTGGCGGTAGCGTTCCTGCGGGGCGCGGGGCGCGGGGCGGCTGCG 180

QY 181 GCGCGGGGCGATGAGGTGCGAGACGCTGTGCGGCGAGCGGTAAAGCGGCGCCCGACCGACCC 240
DB 181 GCGCGGGGCGATGAGGTGCGAGACGCTGTGCGGCGAGCGGTAAAGCGGCGCCCGACCGACCC 240

QY 241 CCGGCGACCCCGGACCCCGCGGTGCGAGCTATGCAAAAGGCGCCCGGCGCTCAGAG 300
DB 241 CCGGCGACCCCGGACCCCGCGGTGCGAGCTATGCAAAAGGCGCCCGGCGCTCAGAG 300

QY 301 CCAAAACCGGCGAGCTTGGGCGCGCGGGGCGGAGCGCGCGCATCTCAGCTTGGGC 360
DB 301 CCAAAACCGGCGAGCTTGGGCGCGCGGGGCGGAGCGCGCGCATCTCAGCTTGGGC 360

QY 361 AACGTGCTCAACTACCTGAGACAGGTACACCGTGGCAGCGCTCTTCTTGACATCCAGCAG 420
DB 361 AACGTGCTCAACTACCTGAGACAGGTACACCGTGGCAGCGCTCTTCTTGACATCCAGCAG 420

QY 421 CACTTTGGGGTCAAGGACCGAGCGCGGCTGCTGAGTCAGTGTTCATCTGTAGCTTC 480
DB 421 CACTTTGGGGTCAAGGACCGAGCGCGGCTGCTGAGTCAGTGTTCATCTGTAGCTTC 480

QY 481 ATGGTGGCTGCCCCCATCTTCGGCTACCTGGGCGACCGCTTCAACAGGAAGGTGATTC 540
DB 481 ATGGTGGCTGCCCCCATCTTCGGCTACCTGGGCGACCGCTTCAACAGGAAGGTGATTC 540

QY 541 AGCTGGGCAATTTCTTCTGGTGGCGCTACCTTCCTCCAGCTCCTCATTTCCCGCAGAG 600
DB 541 AGCTGGGCAATTTCTTCTGGTGGCGCTACCTTCCTCCAGCTCCTCATTTCCCGCAGAG 600

QY 601 TACTTCTGGCTGCTGCTCTGTCGGGGGCTGGTGGGCACTCGGGAGGCGCAGCTACTCC 660
DB 601 TACTTCTGGCTGCTGCTCTGTCGGGGGCTGGTGGGCACTCGGGAGGCGCAGCTACTCC 660

QY 661 ACCATGCGCCCCCACTATCATTTGGGCGACCTTTCACCAAGAACACCGGTACGCTCATGTG 720
DB 661 ACCATGCGCCCCCACTATCATTTGGGCGACCTTTCACCAAGAACACCGGTACGCTCATGTG 720

QY 721 TCCGTCTTCTACTTTCGCCATCCCACTGGGCGAGTGGCTGGGCTACATTTACTGGCTCCAGC 780
DB 721 TCCGTCTTCTACTTTCGCCATCCCACTGGGCGAGTGGCTGGGCTACATTTACTGGCTCCAGC 780

QY 781 GTGAAGCAGGCGAGCGCGGAGACTGGGCACTGGGCACTGGCGGTGTCCCTGTCTCGGCGATG 840
DB 781 GTGAAGCAGGCGAGCGCGGAGACTGGGCACTGGGCACTGGCGGTGTCCCTGTCTCGGCGATG 840

QY 841 ATCAGAGAACACTCATCTCATTTCTGGTCCCGGCCACTAAAGGGGTTCATGCCGACCGAG 900
DB 841 ATCAGAGAACACTCATCTCATTTCTGGTCCCGGCCACTAAAGGGGTTCATGCCGACCGAG 900

QY 901 CTCGGGACCAAGCTCAAGCGCCCGGACCTCATGCGTCCGAGATATGAAGCGCCCTGATTCGA 960
DB 901 CTCGGGACCAAGCTCAAGCGCCCGGACCTCATGCGTCCGAGATATGAAGCGCCCTGATTCGA 960

QY 961 AACCGCAGTACGCTTCTCTCTCCCTGGCCAGCTGGGCTGTCTCTTCGCGCACGGGGGCC 1020
DB 961 AACCGCAGTACGCTTCTCTCTCTCCCTGGCCAGCTGGGCTGTCTCTTCGCGCACGGGGGCC 1020

QY 1021 CTGGGCGATGTGGATCCCGCTCTACCTGCACCGCGCCCAAGTTGTGAGAGACAGCAGAG 1080
DB 1021 CTGGGCGATGTGGATCCCGCTCTACCTGCACCGCGCCCAAGTTGTGAGAGACAGCAGAG 1080

QY 1081 ACGTGCAACAGCCCGCTGTGGGGCCAGGACAGCCTCATCTTTTGGGGCCATCACCTGC 1140
DB 1081 ACGTGCAACAGCCCGCTGTGGGGCCAGGACAGCCTCATCTTTTGGGGCCATCACCTGC 1140

QY 1141 TTTACGGGATTTCTGGGCGTGGTCA CGGGGGCAGGAGCCACGCGCTGGTCCCGCTGAAG 1200
DB 1141 TTTACGGGATTTCTGGGCGTGGTCA CGGGGGCAGGAGCCACGCGCTGGTCCCGCTGAAG 1200

QY 1201 ACCAGCGGGCCGACCCCATCTGGTGTGTCGCTGGGCACTGCTGGGCTCTGCATCTTCATC 1260
DB 1201 ACCAGCGGGCCGACCCCATCTGGTGTGTCGCTGGGCACTGCTGGGCTCTGCATCTTCATC 1260

QY 1261 TGCTGTATCTTCTGGTGGTGGCAAGCAGCATCGTAGGACCTATATCTGTATCTTCGTC 1320
DB 1261 TGCTGTATCTTCTGGTGGTGGCAAGCAGCATCGTAGGACCTATATCTGTATCTTCGTC 1320

QY 1321 GGGGAGACGCTGTGTTTCTTAACTGGGCCCATCACTGACAGACATCTCTATGTAGTGGTC 1380
DB 1321 GGGGAGACGCTGTGTTTCTTAACTGGGCCCATCACTGACAGACATCTCTATGTAGTGGTC 1380

QY 1381 ATCCCAACGGGGCGGCGACCTGGCTGGCCCTTCAGAGCTTCACTCCCACTGCTGGGG 1440
DB 1381 ATCCCAACGGGGCGGCGACCTGGCTGGCCCTTCAGAGCTTCACTCCCACTGCTGGGG 1440

QY 1441 GAGCGGGGAGCGCCCTACCTCATTGGCTTATCTCAGAGCTGATCCGCGCAGACACTAAG 1500
DB 1441 GAGCGGGGAGCGCCCTACCTCATTGGCTTATCTCAGAGCTGATCCGCGCAGACACTAAG 1500

Db 241 GCGGGGGGAAGCGCGGCCATCCTCAGCTTGGGCAACGTGCTCAACTACCTGGACAGGT 300
Qy 386 ACACCGTGGCAGCGGTCTCTTCTGGACATCCAGCAGACATTTTGGGGTCAAGGACCGAGGCG 445
Db 301 ACACCGTGGCAGCGGTCTCTTCTGGACATCCAGCAGACATTTTGGGGTCAAGGACCGAGGCG 360
Qy 446 CCGGCTGTCTGAGTCAAGTGTTCATCTGTAGCTTCATGTGGTGGTGGCCCATCTTCGGGT 505
Db 361 CCGGCTGTCTGAGTCAAGTGTTCATCTGTAGCTTCATGTGGTGGTGGCCCATCTTCGGGT 420
Qy 506 ACCTGGCGCACCGCTTCAACAGGAAGGTATCTCAGCTGCGGCAATTTTCTTCTGTGTGG 565
Db 421 ACCTGGCGCACCGCTTCAACAGGAAGGTATCTCAGCTGCGGCAATTTTCTTCTGTGTGG 480
Qy 566 CCGTCACTCTTCCAGTCTCTTCAATCCCGCAGAGTATCTTGGCTGTGTGGTCTCTCTCC 625
Db 481 CCGTCACTCTTCCAGTCTCTTCAATCCCGCAGAGTATCTTGGCTGTGTGGTCTCTCTCC 540
Qy 626 GGGGGCTGTGGGCATCGGGGAGCGCAGCTACTCCACCATCGCCGCCCATCATTCATTTGGG 585
Db 541 GGGGGCTGTGGGCATCGGGGAGCGCAGCTACTCCACCATCGCCGCCCATCATTCATTTGGG 600
Qy 686 ACCTCTTCAACAAGAACACCGTACGCTCATGTCTCGTCTTCTACTTCTGCGCATCCAC 745
Db 601 ACCTCTTCAACAAGAACACCGTACGCTCATGTCTCGTCTTCTACTTTCGCCATCCAC 560
Qy 746 TGGGCAGTGGCTTGGGCTACATTTACTGGTCCAGCTGAAGCAGGAGCGGAGAGTGGC 805
Db 661 TGGGCAGTGGCTTGGGCTACATTTACTGGCTCCAGCTGAAGCAGGAGCGGAGAGTGGC 720
Qy 806 ACTGGGCATTTGGGGTGTCCCTCTCTGGGCATGATCACGGAACACTCATCTCTCTTC 865
Db 721 ACTGGGCATTTGGGGTGTCCCTCTCTGGGCATGATCACGGAACACTCATCTCTCTTC 780
Qy 866 TGTCTCCAGCACCTAAAGGGGTCTATGCCGACAGCTCGGGGACAGCTCAAGGCCCGGA 925
Db 781 TGTCTCCAGCACCTAAAGGGGTCTATGCCGACAGCTCGGGGACAGCTCAAGGCCCGGA 840
Qy 926 CTTCTATGGTCCGAGATATGAAGGCCCTGATTGAAACCGCAGCTACGTCTTCTCTCC 985
Db 841 CTTCTATGGTCCGAGATATGAAGGCCCTGATTGAAACCGCAGCTACGTCTTCTCTCC 900
Qy 986 TGGCCAGTGGCTGTCTCTTCCGACGCGGGCCCTGGGCAATGTCGATCCCGCTCTACC 1045
Db 901 TGGCCAGTGGCTGTCTCTTCCGACGCGGGCCCTGGGCAATGTCGATCCCGCTCTACC 960
Qy 1046 TGCACCGCGCCCAAGTTGTGCAAGACAGACAGACGTGCAACAGCCCGCCCTGTGGGG 1105
Db 961 TGCACCGCGCCCAAGTTGTGCAAGACAGACAGACGTGCAACAGCCCGCCCTGTGGGG 1020
Qy 1106 CCAAGGACAGCTCATCTTTGGGGCCATCATCTGCTTTTACGGGATTTCTGGGGTGGTCA 1165
Db 1021 CCAAGGACAGCTCATCTTTGGGGCCATCATCTGCTTTTACGGGATTTCTGGGGTGGTCA 1080
Qy 1166 CCGGGGAGAGGACGCGCTGTGCGCTGAAGACCCAGCGGGCGGACCCACTGGTGT 1225
Db 1081 CCGGGGAGAGGACGCGCTGTGCGCTGAAGACCCAGCGGGCGGACCCACTGGTGT 1140
Qy 1226 GTCCGCTGGGCACTGTGGGCTGTGCCATCTTCACTGCTGATCTTCTGGTGGCTGCCAAGA 1285
Db 1141 GTCCGCTGGGCACTGTGGGCTGTGCCATCTTCACTGCTGATCTTCTGGTGGCTGCCAAGA 1200
Qy 1286 GCAGATCTGAGAGCTATATCTGTATCTTCTGGGGAGAGCGTGTGTTTCTAACT 1345
Db 1201 GCAGATCTGAGAGCTATATCTGTATCTTCTGGGGAGAGCGTGTGTTTCTAACT 1260
Qy 1346 GGGCCATCATCTGCAGACATCTCTATGTGTGTATCCCGCGCGGCGGACCTGCGG 1405
Db 1261 GGGCCATCATCTGCAGACATCTCTATGTGTGTATCCCGCGGCGGCGGACCTGCGG 1320
Qy 1406 TGGCTTTGCAGAGCTTTCACCTCCACCTGCTGGGGAGCGCGGAGCCCTACCTCATTTG 1465
Db 1321 TGGCTTTGCAGAGCTTTCACCTCCACCTGCTGGGGAGCGCGGAGCCCTACCTCATTTG 1380

Qy 1466 GCTTATCTCAGACCTGATCCGCCAGAGCACTAAGCACTCCCCGCTCTGGGAGTTCCTGA 1525
Db 1381 GCTTATCTCAGACCTGATCCGCCAGAGCACTAAGCACTCCCCGCTCTGGGAGTTCCTGA 1440
Qy 1526 GCTCGGGGTAGCGCTCATGTCTCTCCCTTTGCTGTGTCTGCGCGCATGTTCTTCC 1585
Db 1441 GCTCGGGGTAGCGCTCATGTCTCTCCCTTTGCTGTGTCTGCGCGCATGTTCTTCC 1500
Qy 1586 TGGCCACTGCGCTCTTCTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACCTG 1638
Db 1501 TGGCCACTGCGCTCTTCTTCTGTCAGCGACCGCGCCAGGGCTGAGCAGCACCTG 1553
RESULT 3
ADE07415
ID ADE07415 standard; DNA; 1250 BP.
XX AC ADE07415;
XX DT 29-JAN-2004 (first entry)
XX DE Novel coding sequence (useful for identifying genetic disorders) #481.
XX KW novel gene; novel protein; tissue marker; molecular weight marker;
XX KW chromosome marker; genetic disorder; gene; ds.
XX OS Unidentified.
XX PN WO2003054152-A2.
XX PD 03-JUL-2003.
XX PF 10-DEC-2002; 2002WO-US039555.
XX PR 10-DEC-2001; 2001US-0339739P.
XX PR 11-DEC-2001; 2001US-0339453P.
XX PR 14-MAR-2002; 2002US-0365091P.
XX PR 14-MAR-2002; 2002US-0365384P.
XX PR 12-APR-2002; 2002US-0372381P.
XX PR 22-APR-2002; 2002US-0372615P.
XX PR 24-APR-2002; 2002US-00128558.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.
DR P-PSDB; ADE08326.
XX PT New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX PS Claim 1; SEQ ID NO 481; 1177bp; English.
XX CC The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present DNA sequence represents a gene of the
CC invention.
XX SQ Sequence 1250 BP; 202 A; 407 C; 419 G; 222 T; 0 U; 0 Other;
Query Match 45.1%; Score 933.4; DB 10; Length 1250;
Best Local Similarity 99.9%; Pred. No. 2.3e-161;
Matches 934; Conservative. 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1564 GTCCTGGCGGCATGTTCTTCTCCGCACTGCGCTCTTCTTCGTACGACCGCGCCAGG 1623
Db |||||
555 GTCCTGGCGGCATGTTCTTCTCCGCACTGCGCTCTTCTTCGTACGACCGCGCCAGG 614
Qy 1624 GCTGAGCAGCAGCTG 1638
Db |||||
615 GCTGAGCAGCAGGTTG 629

RESULT 6

AAI61115

ID AAI61115 standard; cDNA; 1727 BP.

XX AC AAI61115;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 5104.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX KW peripheral nervous system; neuropathy; central nervous system; CNS;

XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;

XX KW Chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

XX KW leukaemia; ss.

XX OS Homo sapiens.

XX PN W0200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;

XX PI Zhou P, Goodrich R, Drmanac RT;

XX DR WPI; 2001-442253/47.

XX DR P-PSDB; AAM41939.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders such

XX PT as central nervous system injuries.

XX PS Claim 1; SEQ ID NO 5104; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the

XX CC encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

XX CC localised neuropathies and central nervous system diseases, such as

XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX CC utilisation of the activities such as: Immune system suppression,

XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

XX CC assays for receptor activity, arthritis and inflammation, leukaemias and

XX CC C.N.S disorders. Note: The sequence data for this patent did not form

CC part of the printed specification

XX SQ Sequence 1727 BP; 306 A; 569 C; 519 G; 332 T; 0 U; 1 Other;
Query Match 29.2%; Score 603.8; DB 4; Length 1727;
Best Local Similarity 98.9%; Pred. No. 4.3e-101;
Matches 608; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1024 GGCATGTGGATCCGCTCTACCTGCACCGCGCCCAAGTTGTGCAAGAGACAGACGACG 1083
Db |||||
15 GGCATGTGGATCCGCTCTACCTGCACCGCGCCCTTGTGTGATAGACAGACGACG 74

Qy 1084 TGCAACAGCCCGCTGTGGGGCAAGGACAGCCTCATCTTTGGGGCCATCACCCTGCTT 1143
Db |||||
75 TGCAACAGCCCGCTGTGGGGCCCAAGGACAGCCTCATCTTTGGGGCCATCACCCTGCTT 134

Qy 1144 ACGGGATTTCTGGGCGTGGTACAGGGGGGAGGAGCCACGCGCTGGTCCGCTGAAGACC 1203
Db |||||
135 ACGGGATTTCTGGGCGTGGACACGCGGGGCGAGAGCCACGCGCTGGTCCGCTGAAGACC 194

Qy 1204 CAGCGGCGCCACCACTGGTGTGCGGTGGGCATGCTGGGCTCTGCCATCTTCATCTGC 1263
Db |||||
195 CAGCGGCGCCACCACTGGTGTGCGGTGGGCATGCTGGGCTCTGCCATCTTCATCTGC 254

Qy 1264 CTGATCTTCGTGGCTGCCAAGACAGCAGCATCGTAGGAGCCTATATCTGTATCTTTCGTCGGG 1323
Db |||||
255 CTGATCTTCGTGGCTGCCAAGACAGCAGCATCGTAGGAGCCTATATCTGTATCTTTCGTCGGG 314

Qy 1324 GAGACGCTGCTTTTCTAACTGGGCGCATCAGTCGAGACATCCTCATGTAGTGGTCAATC 1383
Db |||||
315 GAGACGCTGCTTTTCTAACTGGGCGCATCAGTCGAGACATCCTCATGTAGTGGTCAATC 374

Qy 1384 CCCACGCGCGCCCACTGCGGTGGCTTGCGAGAGCTTACCTCCCACTGCTGGGGGAC 1443
Db |||||
375 CCCACGCGCGCCCACTGCGGTGGCTTGCGAGAGCTTACCTCCCACTGCTGGGGGAC 434

Qy 1444 GCCGGAGCGCCCTACCTCATTTGCTTTATCTCAGACCTGATCGCCAGACACTAAGGAC 1503
Db |||||
435 GCCGGAGCGCCCTACCTCATTTGCTTTATCTCAGACCTGATCGCCAGACACTAAGGAC 494

Qy 1504 TCCCGCTCTGGAGTTCCTGAGCCTGGGCTAGCGCTCATGCTCTGCCCTTTCGTCGTG 1563
Db |||||
495 TCCCGCTCTGGAGTTCCTGAGCCTGGGCTAGCGCTCATGCTCTGCCCTTTCGTCGTG 554

Qy 1564 GTCTCGGGCGCATGTTCTTCCTCGCACCTGCGCTCTTCTTCGTACGAGCAGCGCCAGG 1623
Db |||||
555 GTCTCGGGCGCATGTTCTTCCTCGCACCTGCGCTCTTCTTCGTACGAGCAGCGCCAGG 614

Qy 1624 GCTGAGCAGCAGCTG 1638
Db |||||
615 GCTGAGCAGCAGGTTG 629

RESULT 7

ABQ74263

ID ABQ74263 standard; cDNA; 1587 BP.

XX AC ABQ74263;

XX DT 14-OCT-2002 (first entry)

XX DE Human 46455 transporter protein encoding cDNA SEQ ID NO:6.

XX KW Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;

XX KW nootropic; neuroprotective; hypotensive; antidepressant; neuroleptic;

XX KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;

XX KW obesity; diabetes; anorexia; central nervous system disorder; depression;

XX KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;

XX KW autonomic function disorder; schizophrenia; learning disorder; amnesia;

XX KW memory disorder; age-related disorder; cardiovascular disorder;

XX KW ischaemia reperfusion injury; restenosis; hormonal disorder;

XX KW hypothyroidism; hyperthyroidism; immune disorder;

XX KW chronic mucocutaneous candidiasis; gene; ss.

Db 1326 TGACGGCTGCGCGGAACCTGCCCCCTCTCTTGTGCGAGTTCCGGGCTCTGCAGTT 1385
 Qy 1536 CGCGCTCATGCTCTGCCCCCTTTGCTGCTGCTGCGCGGCATGTTCTTCTCGCACTGCG 1595
 Db 1386 CTCGCTCATGCTCTGCGCGTTGTTGGGCACTGGGCGCGCAGCCTTCTTGGCAGCGC 1445
 Qy 1596 GCTCTTCTGCTGACGACCGCGGCTGAGGCTGAGCAGCAGCTGGGG 1642
 Db 1446 CATCTTCAATTGAGCGCGGCGCGCGGCGGCGACAGCTGCACGTGCAGG 1492

RESULT 8

ADD37479

ID ADD37479 standard; cDNA; 1587 BP.

XX AC

XX AC

XX AC

XX 15-JAN-2004 (first entry)

XX DT

XX DE Human transporter 46455 cDNA #2.

XX KW

XX KW Human; ss; gene; transporter; cytostatic; anorectic; antidiabetic;

XX KW anticonvulsant; gene therapy; PGC-1 associated disorder; liver tumour;

XX KW obesity; epilepsy; diabetes.

XX OS

XX OS Homo sapiens.

XX PN

XX PN US2003143675-A1.

XX PD

XX PD 31-JUL-2003.

XX PF

XX PF 22-MAY-2002; 2002US-00154419.

XX PR

XX PR 12-MAY-2000; 2000US-0204211P.

XX PR

XX PR 29-JUN-2000; 2000US-0215376P.

XX PR

XX PR 31-JUL-2000; 2000US-021769P.

XX PR

XX PR 19-SEP-2000; 2000US-0233790P.

XX PR

XX PR 25-SEP-2000; 2000US-0235107P.

XX PR

XX PR 05-OCT-2000; 2000US-0238336P.

XX PR

XX PR 14-NOV-2000; 2000US-0248364P.

XX PR

XX PR 15-NOV-2000; 2000US-0248878P.

XX PR

XX PR 15-DEC-2000; 2000US-0256240P.

XX PR

XX PR 18-DEC-2000; 2000US-0256588P.

XX PR

XX PR 21-DEC-2000; 2000US-0258028P.

XX PR

XX PR 22-JAN-2001; 2001US-0263169P.

XX PR

XX PR 14-MAY-2001; 2001US-00858194.

XX PR

XX PR 29-JUN-2001; 2001US-00895811.

XX PR

XX PR 31-JUL-2001; 2001US-00919781.

XX PR

XX PR 19-SEP-2001; 2001US-00957664.

XX PR

XX PR 25-SEP-2001; 2001US-00964295.

XX PR

XX PR 05-OCT-2001; 2001US-00972724.

XX PR

XX PR 14-NOV-2001; 2001US-00002769.

XX PR

XX PR 17-DEC-2001; 2001US-00024623.

XX PR

XX PR 22-JAN-2002; 2002US-00055025.

XX PA

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI

XX PI Curtis RAJ, Glucksmann MA, Meyers RE;

XX XX

XX DR WPI; 2003-851783/79.

XX DR P-PSDB; ADD37478.

XX XX

XX PT New isolated nucleic acid, useful for preparing a composition for

XX PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy

XX PT or diabetes.

XX XX

XX PS Claim 1; SEQ ID NO 56; 663pp; English.

XX XX

XX CC The invention relates to an isolated nucleic acid comprising a cDNA

XX CC encoding a human transporter protein, or its complement, a sequence that

XX CC is 60 % identical to the cDNA, a fragment comprising at least 30

XX CC nucleotides of the cDNA, or a sequence encoding a fragment of the

XX CC polypeptide comprising at least 10 contiguous amino acid residues of the

XX CC

CC cDNA. Also included are a vector comprising the novel nucleic acid
 CC molecule, producing the polypeptide, the isolated transporter
 CC polypeptide, an isolated antibody that specifically binds to the
 CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
 CC a sample, a kit, identifying a compound that binds to, or that modulates
 CC the activity of, the polypeptide, and modulating the activity of the
 CC polypeptide. The nucleic acid is useful for preparing a composition for
 CC treating PGC-1 (not defined) associated disorders e.g. liver tumors,
 CC obesity, epilepsy or diabetes. The present sequence encodes a novel human
 CC transporter protein.

XX SQ Sequence 1587 BP; 218 A; 529 C; 479 G; 361 T; 0 U; 0 Other;

Query Match 27.2%; Score 563.8; DB 10; Length 1587;

Best Local Similarity 64.3%; Pred. No. 8.7e-94;

Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

Qy 282 GGGCCCCGGGCTCAGCAGCGCCAAACCGGCCAGCTTGGCGCGGGGGGAGCGCGC 341

Db 126 GGACAGAGGGGCTGCAGCGCATCACCGGCTGTCTCCCGCGCTTCGCTCTCATAGT 185

Qy 342 CGCATCTCTCAGCTTGGGCAACGCTCACTACCTGACAGGTACACCGTGGCAGGGT 401

Db 186 GGGGCTGTGTCTACATCAATCTCTGAACTACATGACCGCTTACCGTGGTGGCGT 245

Qy 402 CCTTCTGACATCCAGCAGCAGCTTTGGGGTCAAGGCGCGAGCGCGCTGTGCAGTC 461

Db 246 CCTTCCGACATCGAGCAGTTCTTCAACATCGGGGACAGTAGCTCTGGCTCATCCAGC 305

Qy 462 AGTGTTCATCTGAGCTTCAATGCTGCTGCCCATCTTGGGCTACCTGGCGAGCGGCT 521

Db 306 CGTGTTCATCTCCAGTTACATGCTGTGGCACTGTCTTGGCTACCTGGTGGACAGGTA 365

Qy 522 CAACAGAGGTGATCTCAGCTGCGGCATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 581

Db 366 CAATCGGAAGTATCTCATGTGCGGGGCTATGCTTCTGCTGCTGCTGCTGCTGCTGCT 425

Qy 582 CTCCTTCATTTCCCGCAGCAGTACTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641

Db 426 ATCTTTCATTTCCCGGAGAGCATTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485

Qy 642 CGGGAGGCGCAGTACTTCCACCATGCGCCCATATCATTTGGGACCTCTTTCACAAAGAA 701

Db 486 CGGGAGGCGCAGTTATTTCCACCATGCGCCCATCTCTCATTTGCCGACCTCTTTGTGGCGGA 545

Qy 702 CAGCGTACGCTCATGCTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 761

Db 546 CCAGCGGAGCGGATGCTCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 605

Qy 762 CTACATTACTTGGCTCCAGCGTGAAGCAGGAGCGGAGAGCTGGGACATTTGGGCTTGGGCT 821

Db 606 CTACATTGAGGCTTCCAAAGTGAAGATATGGCTGGAGACTGGGCTCTGAGGCT 665

Qy 822 GTCCCTGTCTCTGGGATGATCAAGGAACTCATCTCATTTGTGCTCCAGCAGCTAA 881

Db 666 GACACGGGCTAGGAGTGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 725

Qy 882 AAGGGGTATGCGGACGAGCTCGGGA---CGAGCTCAAGCGCGGACCTCATGGCTCCG 938

Db 726 AAGGGGAGCGCTGGAGCGGACCTCAGATTTGGCCACCCCTGAACCCCACTCTGCTGGTGGGC 785

Qy 939 AGATATGAAGCGCTGATTTCGAAACCGAGCTACGCTTCTCTCTCTCTCTCTCTCTCTCTCT 998

Db 786 AGATCTGAGGCTCTTGGCAAGAAATCTAGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 845

Qy 999 TGTCTCTCTCGCACCGGGGCGCTGGGCAATGTGGATCCCGCTCTTACCTGACCGCGGCCA 1058

Db 846 TGTGGCTTTGTCAAGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 905

Qy 1059 AGTTGTGCA---GAAAGACAGCAGAGCTGCAACAGCGCGCGCTGTGGGGCCCAAGGACAG 1115

Db 906 CGTGGTCTTGGGGGAGACCCACCGCTTCCCGGAGACTCTCTGCTCTCTCTCTCTCTCTCT 965

Db 426 ATCTTATCCCGAGAGCATTTCTGGTGTCTCTCTGACCCGGGSCCTGGTGGGGT 485
Qy 642 CGGGAGGCCAGCTACTCCACATCGCCCCCACTATCATTTGGCGACCTTTTACCAAGAA 701
Db 486 CGGGAGGCCAGTTATTCCACATCGGCCACCTCTCATTTGCCGACCTTTTGTGGCGGA 545
Qy 702 CACGGTACGCTCATGCTCGGTCTTCTACTTCCGCAATCCCATCTGGGCACTGGCTGGG 761
Db 546 CCAGCGGAGCCGAGTCTCAGCATCTTCTACTTTGGCCATTCGGTGGGCACTGGTCTGGG 605
Qy 762 CTACATTACTGGCTCCAGCGTGAGCAGCGAGCGAGCTGGCACTGGGCACTGGCGGT 821
Db 606 CTACATTGACGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGT 665
Qy 822 GTCCCCCTGTCTGGGCAATGATCACAGGAACACTCATCTCATTTCTGGTCCAGCACTAA 881
Db 666 GACACCGGCTCTAGAGTGGTGGCGTTCTGCTGCTTCTGTGTAGTGGCGGAGCGGCC 725
Qy 882 AAGGGTCAATCGGACACGCTCGGGGA---CCAGCTAAGGCCCGGACCTCATGGCTCCG 938
Db 726 AAGGGGAGCCGTGGAGCGCCACTCAGATTTCGACCCCTGAACCCCACTCTGTTGGGC 785
Qy 939 AGATATGAAGCCCTGATTTCAACCGCAGCTAGCTTCTCTCTCCCTGGCCACGTCGGC 998
Db 786 AGATCTGAGGCTCTGGCAAGAAATCTAGTTTCTGCTCTGCTTCTCCCTGGGCTTCACTGC 845
Qy 999 TGTCTCTTCCCGACGGGGCCCTGGGCAATGGATCCGCTCTACCTGACCCGCGCCA 1058
Db 846 TGTGGCTTTGTACAGGGCTCCCTGGCTCTGTGGGCTCCGCAATTCCTGCTGGTCCCG 905
Qy 1059 AGTTGTGCA---GAAGACAGCAGAGAGTGCAACAGCCGCCCTGTGGGGCCAAAGACAG 1115
Db 906 CGTGGTCTTGGGAGAGACCCACCTGCTTCCCGAGACTCTCTCTCTCTCTGACAG 965
Qy 1116 CCTCATCTTGGGGCCATCACTCTCTTACGGGATTTCTGGGCTGTGGGCGGAGG 1175
Db 966 TCTCATCTTTGGAATCATCACTCTGCTGACCGGAGTCTGGGCTGTGGGCTGTGGGA 1025
Qy 1176 AGCCACGCGCTGTGGCGCCCTGAGAGACCCAGCGGCGCCACCTGTTGTGTGGCGGG 1235
Db 1026 GATCAGCGCGGCTCGGCCACTCCACCCCGGGCTGATCCCTGTGTGTGGCACTGG 1085
Qy 1236 CATCTGGGCTCTGCCATTTTCATCTGCTGATTTCTGGGCTGCCAAGACAGCATGCT 1295
Db 1086 CCTCTGGGCTCTGACCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145
Qy 1296 AGGAGCTATATCTGTATCTTCTGCTGGGAGACGCTGCTGTTTCTAACTGGGCACTAC 1355
Db 1146 GGCCACTTATATTTTCACTTTCATTTGGAGAGACCTCTCTGTCATGAACCTGGGCACTG 1205
Qy 1356 TGCAGACATCTCATGTACGTGGTTCATCCCAAGCGGCGCCACTGCGGTGGCTTGCA 1415
Db 1206 GGCCGACATTTCTGTACGTGGTGTATCCCTACCGAGCTTCCACCGCGGAGGCTTCCA 1265
Qy 1416 GAGCTTCACTTCCACCTGTCTGGGGGAGCGCGGAGGCCCTTACCTCATTTGCTTTATCTC 1475
Db 1266 GATCGTGTGTCTCCACCTGTCTGGGTGATGCTGGGAGCCCTTACCTCATTTGSCCTGATCTC 1325
Qy 1476 AGACTGATTCGCCAGAGCACTAAGGACTCCCGCTCTGGAGTTCTGAGCCTGGGCTA 1535
Db 1326 TGACCGCTGCGCGGAACTGGCCCCCTCTCTTCTGTCTGAGTTCTGGGCTCTGCACTT 1385
Qy 1536 CGGCTCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1595
Db 1386 CTCCTCATGCTCTGCGGTTTGTGGGCACTGGGCGGCACTGGGCGGCACTGGGCGGCACTGG 1445
Qy 1596 GCTCTTCTCTGACGCAACCGCGGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1642
Db 1446 CATCTTCAATTGAGCGCGGAGCGCGCGGCGGCACTGACGCTGACGCTGACG 1492

RESULT 10
ADL61203

ID ADL61203 standard; DNA; 2165 BP.
XX AC
XX ADL61203;
DT 03-JUN-2004 (first entry)
XX Human protein tyrosine kinase biomarker spinster-like DNA.
DE predictor set; protein tyrosine kinase; cytostatic; antiangiogenic;
XX vasotrophic; vulnery; pharmacogenomic; drug sensitivity; breast cancer;
KW hypervascular disease; angiogenesis; wound healing scar; human;
KW biomarker; ds; gene; spinster-like.
XX Homo sapiens.
OS
XX WO2004020583-A2.
XX 11-MAR-2004.
XX 26-AUG-2003; 2003WO-US026491.
XX 27-AUG-2002; 2002US-0406385P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Huang F, Han X, Reeves KA, Amler L, Fairchild CR, Lee FY;
PI Shaw P;
XX WPI; 2004-239171/22.
DR P-PSDB; ADL61327.
XX
XX New predictor sets with a plurality of polynucleotides and/or
XX polypeptides whose expression pattern predicts cell response to a
XX compound that modulates protein tyrosine kinase activity, useful in
XX treating breast cancer.
XX Claim 2; SEQ ID NO 127; 649pp; English.
XX
XX The invention relates to a novel predictor set comprising a plurality of
XX polynucleotides and/or polypeptides whose expression pattern is
XX predictive of the response of cells to treatment with a compound that
XX modulates protein tyrosine kinase activity or members of the protein
XX tyrosine kinase pathway. The molecules of the invention demonstrate
XX cytostatic, antiangiogenic, vasotrophic and vulnery activities and may
XX be useful in the field of pharmacogenomics, in particular for determining
XX drug sensitivity and in treating breast cancer, hypervascular diseases,
XX angiogenesis and scars in wound healing. The current sequence is that of
XX a human protein tyrosine kinase biomarker DNA of the invention.
SQ Sequence 2165 BP; 308 A; 729 C; 662 G; 466 T; 0 U; 0 Other;

Query Match 27.28; Score 563.8; DB 12; Length 2165;
Best Local Similarity 64.3; Pred. No. 9e-94;
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;
Qy 282 GGGCCCCGGGCTCAGCAGCCAAACCGGCCAGCTTGGCGCGCGGGGGGCGCGC 341
Db 450 GGACGAGAGGGGTGAGCGCATCACCGGCTGTCTCCGGCGCTTGGCTCTCATAGT 509
Qy 342 CGCCATCTCAGCTTGGGCAACGTGTCAACTACCTGACAGGTACACCGTGGAGCGGT 401
Db 510 GCGGCTGTGTCTTACATCAATCTCTGAACTACATGACCGCTTACCGTGGTGGGT 569
Qy 402 CCTTCTGGACATCCAGCAGCATTGGGGTCAAGGACCGAGGCCCGGCTGTGCAGTC 461
Db 570 CCTTCCGACATCGAGCAGTTCTTCAACATCGGGCAGAGTAGTCTCGGGCTCATCCAGAC 629
Qy 462 AGTGTTCATCTGTAGCTTCAATGTGGTGTGCCCCCATCTTGGCTACCTGGGCGACGCTT 521
Db 630 CGGTTCATCTCCAGTTACATGTTGGGACCTGTGTTGGCTTACCTGGTGTACAGGTA 689
Qy 522 CAACAGGAGGTGATTTCTCAGCTGCGGCATTTTCTTCTGTGGCGCGCTCACCTTCTCCAG 581

Db 690 CAATCGAAGATATCTCATGTGCGGGGCAATTGCTTTCTGCTCCCTGGTGACACTGGGGTC 749
Qy 582 CTCCTTCATTTCCCGACGACTTCTGGGTCTGCTGCTCTGTCCTCCGGGGCTGGTGGGCAT 641
Db 750 ATCTTCATCCCGGAGAGCATTTCTTGCTGCTCTCTCTCTGACCCCGGGGCTGGTGGGGT 809
Qy 642 CGGGAGGCGAGCTACTTCCACCATCGCCGCCACTATCATTTGGGAGCTCTTCCACCAAGAA 701
Db 810 CGGGAGGCGAGTTATTCCACCATCGGCCCACTCTCATTTGGGACCTCTTTGTGGCCA 869
Qy 702 CAGCGTAGCGTCATGCTGTCGCTTCTTACTTTCGCAATCCCACTGGGAGTGGCTGGG 761
Db 870 CCAGCGGAGCGGATGCTCAGCATCTTCTACTTTGCAATTCGGTGGGAGTGGTCTGGG 929
Qy 762 CTACATTTAGGCTCAGCGGTGAAGCAGGCGCGGAGCTGGCACTGGGCAATGGGGT 821
Db 930 CTACATTTGAGGCTCAAAAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGT 989
Qy 822 GTCCCTCTCTGGGCAATGATCACAGGAACACTCATCTCTCTGCTCCAGCCACTAA 881
Db 990 GACACCGGCTTAGGAGTGGTGGCCGTTCTGCTGCTGTTCTCTGTTAGTGGGAGCGCC 1049
Qy 882 AAGGGTCTATGCCGACAGCTGGGA---CCAGCTCAAGGCCCGGACCTCATTGGCTCCG 938
Db 1050 AAGGGAGCGCTGGAGCGCACTCAGATTTGCCACCCCTGAACCCCACTCTGCTGGTGGG 1109
Qy 939 AGATATGAAGCCCTGATTCGAACCGCAGCTACGTTCTCTCTCTGCGCCACGCTCGG 998
Db 1110 AGATCTGAGGGCTCTGGCAAGAAATCCTAGTTTGGCTCTGTTCTCTCTGGGCTTCACTGC 1169
Qy 999 TGTCTCTCTCGCCACCGGGGCGCTGGGCAATGGATCCCGCTCTACTGTCACCGCCCA 1058
Db 1170 TGTGGCTTTGTACCGGCTCCTGCTCTGTGGGCTCGGCATTCCTGCTGGTTCGCG 1229
Qy 1059 AGTTGTGCA---GAAGACAGCAGAGACGTGCAACAGCCCGCTGTGGGGCCAGACAG 1115
Db 1230 CGTGGTCTTGGGGAGACCCACCTCTCCCGGAGACTCTCTCTCTCTGACAG 1289
Qy 1116 CTTATCTTTTGGGGCATCATCTGCTTTAGGGATTTCTGGGGCTGTCACGGGGCAGG 1175
Db 1290 TCTCATCTTTTGGACTCATCACCTGCTGACCGGAGTCTTGGGTGGGGCTGGGTGGA 1349
Qy 1176 AGCCAGCGCTGTGCGCTGAAGACCCAGCGGGCGGACCCCACTGCTGTGTGCGCTGGG 1235
Db 1350 GATCAGCGCGGCTCGGCCACTCCAAACCCCGGGCTGATCCCTGCTGTGCTGCTG 1409
Qy 1236 CATGCTGGGCTTGCCATCTTCAATCTGCTGATCTTCTGGGTGCGCAAGAGCAGCATG 1295
Db 1410 CTTCTCTGGGCTGCACT 1469
Qy 1296 AGGAGCTATATCTGTATCTTCTGCGGGAGAGCGTCTGTTTCTTAACTGGGCCATC 1355
Db 1470 GGCCATTTATTTTATCTTTATTTGAGAGACCTCTCTGTCATGAATGGGCCATG 1529
Qy 1356 TGAGACATCTCATGTGCTGATCCCAAGCGGGCGGACCTGCTGCTGCTGCTGCTGCA 1415
Db 1530 GGCGGACATCTGCTGATGCTGATCTCTACCCAGAGCTCAACCGCGGAGGCTTCCA 1589
Qy 1416 GAGCTTCACTCCACTGCTGGGAGCGCGGAGCCCTTACCTCATTTGGCTTTATCTC 1475
Db 1590 GATCGTGTCTCCACCTGCTGGGTGATGCTGGGAGCCCTTACCTCATTTGGCTGATCTC 1649
Qy 1476 AGAGCTGATCCGCGAGACACTAAGAGCTCCCGCTCTGGAGTCTCTGAGCTGGGCTA 1535
Db 1650 TGACCGGCTCGCGGAACTGGCCCCCTCTCTTCTTCTGTCGAGTTCCTGGGCTCTG 1709
Qy 1536 CGGCTCATGCTCTGCGCTTCTGCTGCTGCTGCTGGGCGGATGTTCTCTCTGCCACTGC 1595
Db 1710 CTGCTCATGCTCTGCGCTTCTGCTGGGCACTTGGGGCGGAGCCCTTCTGCGGACCG 1769
Qy 1596 GCTCTTTCTTGTGAGCGACCGCGCCAGGGGTGAGCAGACCTGGGGG 1642
Db 1770 CATCTTTCATTTAGGCGGACCGCGCGGGGACAGCTGCAGCTGCAGG 1816

RESULT 11

AAI57853
ID AAI57853 standard; cDNA; 2218 BP.

XX

AC AAI57853;

XX 22-OCT-2001 (first entry)

DT

XX Human polynucleotide SEQ ID NO 56.

DE

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

XX Homo sapiens.

OS

XX WO200153312-A1.

PN

XX 26-JUL-2001.

PD

XX 26-DEC-2000; 2000WO-US034263.

PF

XX 23-DEC-1999; 99US-00471275.

PR

PR 21-JAN-2000; 2000US-00488725.

PR

PR 25-APR-2000; 2000US-00552317.

PR

PR 20-JUN-2000; 2000US-00598042.

PR

PR 19-JUL-2000; 2000US-00620312.

PR

PR 03-AUG-2000; 2000US-00653450.

PR

PR 14-SEP-2000; 2000US-00662191.

PR

PR 19-OCT-2000; 2000US-00693036.

PR

PR 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

PA

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Zhou J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;

XX WPI: 2001-442253/47.

DR P-PSDB; AAM38697.

XX

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.

PT

XX

XX

PS

PS Claim 1; SEQ ID NO 56; 10078pp; English.

XX

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX

SQ Sequence 2218 BP; 318 A; 745 C; 683 G; 472 T; 0 U; 0 Other;

Query Match 27.2%; Score 563.8; DB 4; Length 2218;

Best Local Similarity 64.3%; Pred. No. 9e-94;

Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

| | | | |
|----|------|--|------|
| Qy | 282 | GGGCCCCGGCGCTCAGCAGCCAAACCGGCAGCTTGGGCGCGCGGCGCGGCGGCGGCGAGCCGC | 341 |
| Db | 497 | GGACAGAGGGGCTGCAGCGCATCACCGGCGTGTCTCCCGGCGCTTCCGGCTCTCATAGT | 556 |
| Qy | 342 | CGCCATCTCAGCTTTGGCACAAGCTGCTCAACTACCTGGACAGGTATACCGTGGCAGGCGT | 401 |
| Db | 557 | GGCGGTGCTGTGCTACATCATCTCCTGAACTACATGACCGCTTACCGTGGCTTGGCGT | 616 |
| Qy | 402 | CCTTCCTGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGGCGCGCGCTGTGCAGTC | 461 |
| Db | 617 | CCTTCCCGACATCGAGCAGTTCTTCAAATCGGGGACAGTAGTCTCGGGCTCATCCAGAC | 676 |
| Qy | 462 | AGTGTTCATCTGTAGCTTCACTGCTGGCTGCCCGCCATCTTCGGCTACCTGGCGGACCGCTT | 521 |
| Db | 677 | CGTGTTCATCTCCAGTTACATGGGTGTGGCACCCTGTGTTTGGCTACCTGGGTGACAGGTA | 736 |
| Qy | 522 | CAACAGGAAGGTGATTCACGCTGGCGGATTTCTTCCTGTGTCGGCGCTCACCTTCTCCAG | 581 |
| Db | 737 | CAATCGGAAAGTATCTCATGTGCGGGGGATTTGCCCTTCTGGTCCCTGTGTGACACTTGGGTC | 796 |
| Qy | 582 | CTCCTTCAATCCCGACGAGTACTTCTTGGCTGCTGCTCCTGTCCCGGGGCTGTGTGGGCAT | 641 |
| Db | 797 | ATCCTTTCATCCCGGAGAGCATTTCTGCTGCTCCTCTGACCCGGGCGCTTGGTGGGGT | 856 |
| Qy | 642 | CGGGAGGCCAGCTACTCCACATCGCCCCCACTATCATTTGGCGACCTCTTTCAACAAGAA | 701 |
| Db | 857 | CGGGAGGCCAGTTATTCCACATCGCGCCCCACTCTATTGCCGACCTCTTTGTGGGCCGA | 916 |
| Qy | 702 | CACGCTAGGCTCATGCTGTCGGTCTTCTACTTTCGGCATCCCATGSGGACGTGGCCCTGGG | 761 |
| Db | 917 | CCAGCGGAGCCGGAATGCTCAGCATCTTCTACTTTTGGCCATTCGGGTGGGACGTGTGTCTG | 976 |
| Qy | 762 | CTACATTACTGGCTCCAGCGGTGAAGACGGCAGCGGAGACTGGCACTGGGGCATTTGCCGGT | 821 |
| Db | 977 | CTACATTGCAGGCTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGGCTCTGAGGGT | 1036 |
| Qy | 822 | GTCCCTGTCTCGGCGATGATCA CAGGAACA CTCACTCTCAATCTGTGTC CAGCCACTAA | 881 |
| Db | 1037 | GACACCGGGTCTAGAGAGTGTGGCGTGTCTGCTGCTGTTCTGTGTAGTTCGGGAGCGCC | 1096 |
| Qy | 882 | AAGGGGTATCCGACACGACTCGGGGA---CCAGCTCAAGGCCCGGACCTCATGGCTCCG | 938 |
| Db | 1097 | AAGGGGAGCCGTGAGGCGCCACTCAGATTTGCCACCCCTGAACCCCACTCGTGGTGGGC | 1156 |
| Qy | 939 | AGATATGAAGCCCTGATTTGAAACCGCAGCTAGTCTTCTTCCTCCCTGGCCACGTCGCGC | 998 |
| Db | 1157 | AGATCTGAGGGCTCTGGCAGAAATCTTAGTTTCTGCTGCTCTTCCCTGGCTTCACTGC | 1216 |
| Qy | 999 | TGCTCTCTTCGCCACGGGGGCCCTGGGCAATGTGATCCCGCTCTACTGTCA CCGCGCCA | 1058 |
| Db | 1217 | TGTGGCCTTTGTACAGGGCTCCCTGGCTCTGTGGGCTCGGCA T TCTGTGCTGCTCCG | 1276 |
| Qy | 1059 | AGTTGTGCA---GAAGACAGCAGAGACGTGCAA CAGCCCGCCCTGTGGGGCCAGGACAG | 1115 |
| Db | 1277 | CGTGGTCTTTGGGGAGACCCACCCCTGCCTTCCCGGAGACTCTCTGCTCTTCTCTGACAG | 1336 |
| Qy | 1116 | CCTCATCTTTGGGGCATCACTGCTTTTACGGGATTTCTGGGCGTGTGTCA CCGGGG CAGG | 1175 |
| Db | 1337 | TCTCATCTTTGAGACTCATCACTGCCTGACCGGAGTCTGGGTGTGGGCTTGGGTGTGGA | 1396 |
| Qy | 1176 | AGCACCGGCTGTGTGCCCTGGAAGACCCAGCGGGCCGACCCACTGGTGTGTGCCGTGGG | 1235 |
| Db | 1397 | GATCAGCCGCGGCTCCGCCACTCCAAACCCCGGGCTGATCCCTGGTCTGTGCCACTG | 1456 |
| Qy | 1236 | CATGCTGGGCTCTGCCATCTTCATCTGGCTGATCTTGTGGCTGCCAAGACGACGATCGT | 1295 |
| Db | 1457 | CCTCCTGGGCTCTGACCCCTTCTCTTCTGCTGCTTGTGCTTGTGCTGCGCCGCTGTGTAGCAT | 1516 |
| Qy | 1296 | AGGAGCCTATATCTGATCTTCTGTCGGGAGACGCTGCTGTTTCTTAAC TGGG C C ATCA C | 1355 |
| Db | 1517 | GGCCACTTATTTTTCATCTTCA TTGAGAGAGCCCTCTCTGTGCTTGTGCTTGTGAGCATCGT | 1576 |
| Qy | 1356 | TGCAGACATCCTCATGTACGTGTGTCA TCCCA CCGCGCGGGCCACTGCCGTGGCTTGCA | 1415 |

RESULT 12

ABQ74262

ID ABQ74262 standard; cDNA; 2230 BP.

AC ABQ74262;

DT 14-OCT-2002 (first entry)

Human 46455 transporter protein encoding cDNA SEQ ID NO:4.

Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian; cardiovascular; neuroprotective; hypotensive; antidepressant; neuroleptic; cardiovascular; immunosuppressive; gene therapy; sugar homeostasis; obesity; diabetes; anorexia; central nervous system disorder; depression; CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension; autonomic function disorder; schizophrenia; learning disorder; amnesia; memory disorder; age-related disorder; cardiovascular disorder; ischaemia reperfusion injury; stenosis; hormonal disorder; hypothyroidism; hyperthyroidism; immune disorder; chronic mucocutaneous candidiasis; gene; ss.

OS Homo sapiens.

| Key | Location/Qualifiers |
|-----|---------------------|
| FH | |
| 200 | |

| FT | key | Education/ |
|----|-----|------------|
| FT | CDS | 376. .1962 |

```
FT
000
/*tag= a
0001.1200
```

FT /product= "46455 transporter protein"

PN WO200255701-A2.

18-JUL-2002.

AA
PF 17-DEC-2001: 2001WO-US049060.XX
PR 15-DEC-2000: 2000US-0256240P.

PR 15-DEC-2000; 2000US-0256240P.
PR 18-DEC-2000; 2000US-0256588P.

FR 19-DEC-2000; 2000US-VZ38388F;
PR 21-DEC-2000; 2000US-0258028P;

PA (MILL-) MILLENNIUM PHARM INC.

PI Curtis RAJ;

XX
DR
WPT: 2002-590672/63.

DR WPI; 2002-390672/
DR P-PSDB; ABP52155.

PT New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins,
PT useful in e.g. treating disorders characterized by insufficient or
PT excessive production of the polypeptides, e.g. cardiovascular or immune
PT disorders.

PS · Claim 1: Fig 8: 364pp: English.

XX CC The present sequence encodes the human 46455 protein, which is a member

CC of the transporter family. The transporter family proteins (1) of the
CC present invention have anorectic, antidiabetic, anti-Parkinsonian,
CC noreptropic, neuroprotective, hypotensive, antidepressant, neuroleptic,
CC cardiovascular and immunosuppressive activities, and can be used in gene
CC therapy. The human transporter proteins from the present invention are
CC designated 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and
CC 67084alt. The polynucleotides encoding (1) can be used in screening
CC assays (e.g. chromosome mapping, tissue typing, or in forensic biology),
CC predictive medicine (e.g. diagnostic assays, prognostic assays,
CC monitoring clinical trials, or pharmacogenetics), as surrogate markers,
CC and in methods of treatment (e.g. therapeutic or prophylactic). (1) are
CC useful for treating disorders characterised by insufficient or excessive
CC production of 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or
CC 67084alt substrates or production of transport 8099, 46455, 54414, 53763,
CC 67076, 67102, 44181, 67084FL or 67084alt inhibitors. Disorders associated
CC with 8099 and 46455 include disorders associated with sugar homeostasis
CC such as obesity, diabetes or anorexia. Disorders associated with 54414,
CC 53763, 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders
CC (e.g. Parkinson's disease or Alzheimer's disease), autonomic function
CC disorders (e.g. hypertension, depression or schizophrenia), or learning
CC or memory disorders (e.g. amnesia or age-related disorders),
CC cardiovascular disorders (e.g. ischaemia reperfusion injury or
CC restenosis), hormonal disorders (e.g. hypothyroidism or hyperthyroidism),
CC or immune disorders (e.g. chronic mucocutaneous candidiasis)

XX
SQ Sequence 2230 BP; 318 A; 752 C; 688 G; 472 T; 0 U; 0 Other;

Query Match 27.2%; Score 563.8; DB 6; Length 2230;

Best Local Similarity 64.3%; Pred. No. 9e-94;

Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

QY 282 GGGCCCCGGCTCAGACGCCCAACCGCCAGCTTGGCCGGGGGGGGGACCCG 341
DB 501 GGACCGAGAGGGCTCGACGCGATCACCGGCTGTCTCCCGCCGTTCCGCTCTCATAGT 560
QY 342 CGCATCTCTCAGCTTGGGCAACGCTGCTCACTACCTGGACAGGTACACCGTGGCAGCGCT 401
DB 561 GGGGGTGTCTGCTACATCACTCTCTGAACTACATGGACCGCTTCAACGCTGGCGCT 620
QY 402 CTTTCTGGACATCCAGACGACATTTGGGGTCAAGGACCGAGGCGCGCGCTGTGAGTTC 461
DB 621 CTTTCCCGACATCGACGAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTATCCAGAC 680
QY 462 AGTGTTCATCTGAGCTTCATGTGTGCTGCCCATCTTCGCTACCTGGGACCGCTT 521
DB 681 CGTGTTCATCTCAGTTACATGTGTGTGGACCTGTGTGGTACCTGGTGACAGTTA 740
QY 522 CAACAGGAAGGTGATTTCTCAGCTGCGGCAATTTTCTTGTGTGCGCGCTCACCTTCTCCAG 581
DB 741 CAATCGGAGTATCTCATGTGCGGGGCAATGCTTCTGTCTCTGTGACACTGGGGTC 800
QY 582 CTCCTTCATTTCCGAGAGTACTTCTGGCTGTGTGCTCTGCTCCGGGGGCTGTGGGCT 641
DB 801 ATCTTTCATCTCCCGGAGAGATTTCTGGCTGTCTCTCTGACCGCGGCGCTGTGGGGT 860
QY 642 CGGGAGGCGGAGTACTTCCACCATCGCCCGGCTATCATTTGCGGAGCTCTTTCACCAAGAA 701
DB 861 CGGGAGGCGGAGTATTTCCACCATCGCGGCGGCTCTCATTTGCCGACCTCTTTGTGGCGA 920
QY 702 CAGCGGTACGCTCATGTGTCTCTTCTACTTTGCGCATCCCACTGGGCAAGTGGCTGGG 761
DB 921 CCAGCGGAGCGGATGCTCAGCATCTTCTACTTTGCAATTCGCTGGGAGTGTCTGGG 980
QY 762 CTACATTTAGGCTTCAGCGTGAAGCGGAGCGGAGACTGGCACTGGGCTTGGGGT 821
DB 981 CTACATTTGAGGCTTCCAAAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGT 1040
QY 822 GTCCCTGTCTCGGCGATGATACAGGAACATCATCTCATTTCTGTCTGCCACCACTAA 881
DB 1041 GACACGGGTCTAGGAGTGTGGCGCTTCTGTCTGTCTGTCTGTGTGTGGGAGCGGCC 1100
QY 882 AAGGGGTCTATGCCGACAGCTCGGGGA---CCAGCTCAAGGCCCGGACCTCATGTGCTCG 938
|||||

DB 1101 AAGGGAGCGCTGGAGCGCCACTCAGATTTGCAACCCCTGAAACCCCACTCTGTGTGGGC 1160
QY 939 AGATATGAAGCCCTGATTCGAAACCGCAGCTACCTCTTCTCTCCCTGGGCCACGCTCGGC 998
DB 1161 AGATCTGAGGCTCTGGCAAGAAATCTAGTTCCTCTCTTCTCCCTGGGCTTACTGC 1220
QY 999 TGTCTCTTTCGCAACGGGGCCCTGGGATGTGGATCCCGCTCTACTCTGACCCGCGCCA 1058
DB 1221 TGTGGCTTTGTACGGGCTCCCTGGCTCTGGGCTCGGCAATCTCTGCTGCTTCCG 1280
QY 1059 AGTTGTGCA---GAAGACAGACAGACGTGCAACAGCCGCCCTGTGGGGCCCAAGACAG 1115
DB 1281 CGTGTCTCTGGGAGAGACCCCACTGCTTCCGAGAGACTCTCTCTCTCTCTGACAG 1340
QY 1116 CTTGATCTTTGGGGCCATCACTGCTTTACGGGATTTCTGGGCGTGGTCAACGGGGCAGG 1175
DB 1341 TCTCATCTTTGAGCTCATACCTGCTGACCGGAGTCTTGGGTGTGGGCTTGGGTGGA 1400
QY 1176 AGCCACGGCTGTGTCGCCCTGAAAGACCCAGCGGCGCCACCACTGGTGTGTGCCGTGG 1235
DB 1401 GATCAGCGCGGCTCCGCCACTCCAAACCCCGGGCTGATCCCTGGTCTGTGCCACTGG 1460
QY 1236 CATGTGGGCTGTGCAATCTTCACTGCTGATCTTCTGTGGGTGCTGCAAGACAGCATCGT 1295
DB 1461 CTTCTGTGGCTGTGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1520
QY 1296 AGGACCTATATCTGATCTTCTGTCGGGAGACGCTGCTGTTTCTTAACCTGGGCCATCAC 1355
DB 1521 GGCCACTTATATTTTCTATCTTCTATTTGGAGAGACCTCTCTGTCATGNACTGGGCCATCGT 1580
QY 1356 TGCAGACATCTCTATGATGCTGATCCCAACCGGCGCGCCACCTGCGCTGGCTTGA 1415
DB 1581 GGCACGACATCTGCTGTAGTGTGATCCCTACCCGAGCTCCACCGCGAGGCTTCCA 1640
QY 1416 GAGCTTACCTCCACTGCTGGGAGACCGCGGAGCCCTACCTCATTTGGCTTTATCTC 1475
DB 1641 GATCGTGTGTCCTCCACTGCTGGGTGATGCTGGGAGCCCTACCTCATTTGGCTGATCTC 1700
QY 1476 AGACCTGATCCCGCAGACACTAAGACCTCCCGCTCTGGGAGTTCTCTGAGCCTGGGCTA 1535
DB 1701 TGACCGCTTGGCGCGGAACCTGGCCCTCTCTCTTGTCCGAGTTCCGGGCTTGAGTT 1760
QY 1536 CGCGCTCATGCTCTGCCCTTTCTGCTGTGTCTGTGGCGGCAATGTTCTTCTCTGCCACTGC 1595
DB 1761 CTCGCTCATGCTCTCGCGCTTTTGGGGCACTTGGGCGCGCAGACCTTCTCTGGCACCCG 1820
QY 1596 GCTCTTCTGTCAGGACCGCGCCAGGCTGAGCAGACCTGGGG 1642
DB 1821 CATCTTCAATTGAGGCGGACCGCGCGGGGCAAGCTGACGCTGCAGG 1867

RESULT 13

ADD37477

ID ADD37477 standard; cDNA; 2230 BP.

XX

AC ADD37477;

XX

DT 15-JAN-2004 (first entry)

XX

DE Human transporter 46455 cDNA #1.

XX

KW Human; ss; gene; transporter; cytostatic; anorectic; antidiabetic;

KW anticonvulsant; gene therapy; PGC-1 associated disorder; liver tumour;

KW obesity; epilepsy; diabetes.

XX

OS Homo sapiens.

XX

PN US2003143675-A1.

XX

PD 31-JUL-2003.

XX

PF 22-MAY-2002; 2002US-00154419.

XX

| | | |
|---------|---------------------|------------|
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,894 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,911 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,636 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,874 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,910 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,864 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,631 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,845 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,892 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/057,761 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/047,595 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,599 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,588 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,585 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,586 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,590 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,594 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,589 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,593 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/047,614 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/043,578 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/043,576 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/047,501 |
| EARLIER | FILING DATE: | 1997-05-23 |
| EARLIER | APPLICATION NUMBER: | 60/043,670 |
| EARLIER | FILING DATE: | 1997-04-11 |
| EARLIER | APPLICATION NUMBER: | 60/056,632 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,864 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,876 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,881 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,909 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,875 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/056,862 |
| EARLIER | FILING DATE: | 1997-08-22 |
| EARLIER | APPLICATION NUMBER: | 60/048,964 |
| EARLIER | FILING DATE: | 1997-06-06 |
| EARLIER | APPLICATION NUMBER: | 60/057,650 |
| EARLIER | FILING DATE: | 1997-09-05 |
| EARLIER | APPLICATION NUMBER: | 60/056,884 |
| EARLIER | FILING DATE: | 1997-08-22 |

```

; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

```

Query Match 26.2%; Score 542.8; DB 4; Length 2201;
Best Local Similarity 64.3%; Pred. No. 7.3e-101;
Matches 877; Conservative 1; Mismatches 478; Indels 8;

| | | | |
|----|------|--|------|
| QY | 285 | CCCCGGCGCTCAGCAGCCGAAACCGGCCGAGCTTGGGCGCGGGCGGGGGGCGAGCCGCGCG | 344 |
| Db | 416 | CCAGGARGGGCTGCAGCGGCATCAGCGCGCTGTCTCTCCGGCGGTTCGGCTCTCATAGTGGC | 475 |
| QY | 345 | CATCCTCAGCTTGGGCAACGTGTCTCACTACTTGGACAGGTACACCTGGCAGGCGTCT | 404 |
| Db | 476 | GGTGCTGTGTACATCAATCTCTTGAATACATGGACCGCTTCAACCTGGCTGGCGTCTCT | 535 |
| QY | 405 | TCTGGACATCCAGCAGCACTTTGGGTTCAGGACCGAGGGCGCGGCTGTGTGACGTACGT | 464 |
| Db | 536 | TCCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGTCTTGGGCTCATCCAGACGT | 595 |
| QY | 465 | GTTTCATCTGTAGCTTCAATGTGGTGGCGCCCATCTTGGGTACTCTGGGCGACCGCTTCAA | 524 |
| Db | 596 | GTTTCATCTCCAGTTACATGTGTGGACCTGTGTGGTCTACTCTGGTGACAGGTACAA | 655 |
| QY | 525 | CAGGAAGTGATCTCAGCTGTGGGCATTTTCTGTGTGGGCCGTCACTTCTCCAGCTC | 584 |
| Db | 656 | TCGGAAGTATCTCATGTGTGGGGGCATTTGCTTCTGGTCCCTGTGTGACACTGGGGTCA | 715 |
| QY | 585 | CTTCATTTCCACAGCAGTACTCTTGGCTGTGTCTCTCGCGGGGCTGTGGGCATCGG | 644 |
| Db | 716 | CTTCATCCCCGGAGAGCATTTCTGGCTGTCTCTCTGACCGGGGCTGTGGGGGTCTGG | 775 |
| QY | 645 | GGAGGCAGCTACTCCACCATCGCCCCCACTATCATTTGGGACCTCTTCAACGAACAC | 704 |
| Db | 776 | GGAGGCAGTTATTCCACCATCGGCCCATCTCTCATTTGCGACCTCTTTGTGGCCGACCA | 835 |
| QY | 705 | GGGTAGCCTCATGTGTCCGTCTTCTACTTTGGCCATCCACTGGGCAAGTGGCCTGGGCTA | 764 |
| Db | 836 | GCGGAC-CGGATGCTCAGCATCTTCTACTTTGCCATTCGGTGGGCAGTGGTCTGGGCTA | 894 |
| QY | 765 | CATTACTGGCTCCAGGTGAAGCAGGACGCGGAGACTGCGCATTTGGGCATTTGGGGTGT | 824 |
| Db | 895 | CATTGCAAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCACTTGGGCTCTGAGGGT | 954 |
| QY | 825 | CCCTGTCTCGGCATGATCACAGGAACATCATCTCATTTGTGTCCCAAGCCCACTAAAG | 884 |
| Db | 955 | ACCGGTCTAGGAGTGTGGCCGTCTGTGTCTGTCTCTGGTAGTGGGAGCGCCCAAG | 1014 |
| QY | 885 | GGGTCTATGCCGACAGCTCGGGGA---CCAGACTCAAGCGCCGACCTCATGGTCTCGAGA | 941 |
| Db | 1015 | GGGAGCGGTGGAGGCCACTCAGATTTGCCACCCCTGAAACCCACCTCTGTGTGGGCAGA | 1074 |
| QY | 942 | TATGAAGCCCTGATTCGAAACCGCACTACGTCTTCTCTCCTGCCACAGTCTGGCTGT | 1001 |
| Db | 1075 | TCTGAGGGCTCTGGCAAGAAATCTAGTTTGGTCTGTCTCTCTGCTGGCTTCACTGTCTGT | 1134 |
| QY | 1002 | CTCTCTTGGCCACGGGGGCTTGGGCATGTGGATCCCGCTCTACTTGCACCGCGCCCAAGT | 1061 |
| Db | 1135 | GGCTTTTGTACCGGGCTCCTGTGCTCTGTGGGCTCCGGCATTCCTGTGCTGTCCCGCT | 1194 |
| QY | 1062 | TGTGCA---GAAGACAGCAGAGAGCTGCAACAGCCCGCTCTGTGGGCGCAAGGACAGCT | 1118 |
| Db | 1195 | GGTCTCTGGGAGACCCCACTTCCGCTTCCGGAGACTCTCTGTCTTCTCTGACAGTCT | 1254 |
| QY | 1119 | CATCTTTGGGGCCATCAGCTGTATTAGGATTTCTGGGCGTGTGTGTACCGGGGCGAGGAGC | 1178 |
| Db | 1255 | CATCTTTGGACTCATCACCTGCTGACCGGAGTCTCTGGGTGTGGGCTGGGTGTGGAGAT | 1314 |
| QY | 1179 | CAGCGCTGTGTGGCTGTGAAGACCAAGCGGGCGGACCCACTGTGGTGTGTGCGTGTGGCAT | 1238 |

[illegible]

Query Match 24.7%; Score 512.4; DB 4; Length 1661;
Best Local Similarity 65.1%; Pred. No. 9.5e-95;
Matches 813; Conservative 3; Mismatches 425; Indels 8; Gaps 4;
QY 400 GTCCTTTCGACATCCAGACGACTTTGGGGTCAAGACCGGCGGCTCTGTCAG 459
DB 1 GTCCTTCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAG 60
QY 460 TCAGTGTTCATCTAGTTCATGGTGGCTGCCCCCATCTTCGGCTACCTGGGGCGACGC 519
DB 61 ACCGTGTTCATCTCCAGTTACATGGTGTGGACCTCTGTGTTGGCTACCTGGGTGACAGG 120
QY 520 TTCAACAGGAGGTGATTCTCAGCTGCGGCAATTTCTTCTGGTGGCGCGCTCACCTCTCC 579
DB 121 TACNATCGGAAGTATCTCATGTGCGGGGCAATGCCCTTGTGCTCCCTGGTGACACTGGG 180
QY 580 AGCTCTTCATTCCTCCAGCAGTACTTCTGGCTGCTGCTCTGCTCCCGGGGCTGGTGGGC 639
DB 181 TCATCTTCATCCCGGAGAGCAATTTCTGGCTGCTCTCTCTGACCCCGGGGCTGGTGGG 240
QY 640 ATCGGGAGGCGACTACTCACCATCGCCCCCACTATCATTTGGCGACCTCTTCCACCAAG 699
DB 241 GTCGGGAGGCGAGTTATTCACCATCGCGCCACTCTCATTTGCGGACCTCTTTGTGGCC 300
QY 700 AACACGGGTAGCTGCTGCTGCTGCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
DB 301 GACACAGGGA--SCGATGCTCAGCATCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG 359
QY 760 GGCTACATTAATGCTCCAGCGTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 819
DB 360 GGCTACATTCAGGCTCCAAAGTGAAGATATGCTGAGACTGGCACTGGGCTCTGAGG 419
QY 820 GTGTCCCTGCTCTGGGCGATGATCAGAGAACATCATCTCTCATTTCTGCTGCTGCTGCTG 879
DB 420 GTGACACCGGCTCTAGGAGTGGTGGCGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 479
QY 880 AAAAGGGGTCTATGCGGACGCTCGGGGA--CCAGCTCAAGGCGCGGACCTCATGGCTC 936
DB 480 CCAAGGGGAGCGGTGAGCGGCACTCAGATTGTCACCCCTGAAACCCCACTCTGTTGGT 539
QY 937 CGAGATATGAAGGCGCTGATTCGAAACCGCAGCTACGTTCTCTCTCTCTCTCTCTCTCTCT 996
DB 540 GCAGATYTAGGGCTCTGGCAAGAAATCTTAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 599
QY 997 GCTGTCTCTTTCGACGCGGCGCTGGGCGATGFGATTCGCTCTACCTGCAACCGGCG 1056
DB 600 GCTGTGGCTTTGTACGCGGCTCCCTGGCTCTGTTGGCTCTGGGCTCTGCTGCTGCTGCTG 659
QY 1057 CAAGTTGTGC--AGAGACAGCAGACAGCTGCAACAGCGCGCTCTGGGGCCAGGAC 1113
DB 660 CGGCTGTCTCTGGGGAGACCCCACTGCTCTCCCGGAGACTCTGCTCTCTCTCTGAC 719
QY 1114 AGCTCATCTTTGGGCGCATCAGCTGTTTACGGGATTTCTGGGCGGTGTCAGGGGGCA 1173
DB 720 AGTCTCATCTTTGGAATCATCACTGCTGACCGGAGTCTGGGTGTGGGCTCTGGGTGTG 779
QY 1174 GGAGGCCACGCTGGTGGTCCGCTGAAAGACCAAGCGGGCGCAACCACTGCTGCTGCTGCTG 1233

DB 780 GAGATCAGCGCGCGYTCGGCCACTCCAAACCCCGGGCTGATCCCTCGTGTCTGTGCCACT 839
QY 1234 GGCATGTGGGCTCTGCATCTTTCATCTGCTGCTGATCTTCTGTTGGCTGCCAAGCAGCATC 1293
DB 840 GGCCTCTGGGCTCTGCACCCCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
QY 1294 GTAGAGCCTATATCTCTATCTTCTGCTCGGGAGACGCTGCTGCTGCTGCTGCTGCTGCTG 1353
DB 900 GTGCCACTTATATTTTTCATCTTCAATTGGAGAGACCTCTCTGCTCCATGAACTGGGCCATC 959
QY 1354 ACTGCAGACATCTCATGTAGTGGTTCATCCCAAGGGCGGGCGGCACTGCGGTGGCTTGG 1413
DB 960 GTGCCCGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
QY 1414 CAGAGCTTCCACCTCCACCTGCTGGGGAGCGCGGGAGCGCCCTACCTCATTTGGCTTTATC 1473
DB 1020 CAGATCGTGTCTGCCACCTGCTGGGTGATGCTGGAGAGCCCTACCTCATTTGGCTGATC 1079
QY 1474 TCAGACCTGATTCGGCCAGACACTTAAGGACTCCCGCTCTGGGAGTTCTTGAGCTGGGC 1533
DB 1080 TCTGACCGCTGCGCGGAACTGGCCCCCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1139
QY 1534 TAGCGCTCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593
DB 1140 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1198
QY 1594 GCGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1642
DB 1199 GNCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1247

RESULT 3
US-09-149-476-87
; Sequence 87, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; EARLIER FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 21.9%; Score 453; DB 4; Length 1460;

Best Local Similarity 65.2%; Pred. No. 9,4e-83;
Matches 716; Conservative 0; Mismatches 375; Indels 8; Gaps 3;

550 ATTTCTTCTGGTGGCGCTCACCTTCTCCAGCTCTTCTCAATTCCTCCAGCAGTACTTCTGG 609
Db 1 ATTGCTTCTGGTCCCTGGTGACACTGGGGTCTATCTTCTATCCCGGAGAGCAATTTCTGG 60
610 CTGCTGTCTCTGTCGGGGGCTGGTGGGCATCGGGAGGCCAGCTACTCCACCATGCC 669
Db 61 CTGCTCTCTCTGACCGGGGCTGGTGGGGTCTGGGGAGGCCAGTATTCCACCATCGCG 120
670 CCCACTATCATGTGGCGACTCTTCCACCAAGAACACGGCTACGCTCATGCTCTCGTCTTC 729
Db 121 CCCACTCTCATGTGGCGACTCTTGTGGCCGACAGCGGACCGGATGCTCAGCATCTTC 180
730 TACTTGGCCATCCCACTGGGCGAGTGGCTGGGCTACATTAATCTGCTCCAGCGTCAAGCAG 789
Db 181 TACTTGGCAATTCGGTGGGCGAGTGTCTGGGCTACATTCAGGCTCCAAAGTGAAGAT 240
790 GCACCGGAGACTGGGCTGGGCAATGGGGTGTCCCTGTCTCTGGGAGATGATCAGAGGA 849
Db 241 ATGGCTGGAGACTGGGCTGGGCTCTGAGGGTGACACCGGGTCTAGGAGTGGTGGCGGTT 300
850 ACATCATCTCTCTGTCGCCAGCCACTAAAGGGGTCTATCCGACCGAGCTCGGGGA- 908
Db 301 CTGCTGTCTCTGTTAGTGGGAGCGCCGCAAGGGAGCGGTGGAGCGCCACTCAGAT 360
909 --CCAGCTCAAGGCCCGGACCTCATGCTCGAGATATGAAGGCCCTGATTCGAAACCGC 966
Db 361 TTGCCACCCCTGAACCCACCTGTGTGGGCGAGATCTGAGGGCTCTGGCAAGAAATCCT 420
967 AGCTAGCTCTTCTCTCTGCGGCACGCTCGGCTGTCTCTTCTGCGCACGGGGGCTTGGGC 1026
Db 421 AGTTTCTGCTGTCTTCTCTGGGCTTCACTGCTGTGGGCTTGTGACGGGCTCTCTGGCT 480
1027 ATGTGGATCCGCTCTACCTGACCGCGCCCAAGTTGTC---AGAGACAGCAGAGACG 1083
Db 481 CTGTGGGCTCGGCAATTCCTGCTGCTTCCGGTTCGGGTCCTTGGGAGACCCCACTCGC 540
1084 TGCAACAGCCCGCTCTGGGGCCCAAGCAGACCTCATCTTTGGGGCCATCACTGTGTTT 1143
Db 541 CTTCCCGGAGACTCTCTCTCTCTGACAGTCTCATCTTGGACTCATCACTCTGCTG 600
1144 ACGGGATTTCTGGCGTGGTCAOCCGGGGCAGGAGCCACGCGCTGGTGGCCCTGAAGACC 1203
Db 601 ACCGGAGTCTGGGTGGGCTGGGTGTGGAGATCAGCGCGCGGCTCCGCCACTCCAAC 660
1204 CAGGGGCCACCCACTGTGTGTGGCTGGGCATGCTGGGCTGCGCATCTTCATCTTC 1263
Db 661 CCGGGGCTGATCCCTGGTGTGGCCACTGGGCTCTGGGCTCTGACCCCTTCTCTTC 720
1264 CTGATCTTCTGGCTGCAAGAGCAGCATCTGAGGACCTATATCTGATCTCTGCTGGG 1323
Db 721 CTGTCCCTTGGCTGGCGCGGTGAGCATGTGGGCACTTATATTTCACTTCATTTGA 780
1324 GAGACGCTGTCTTTTCTAACTGGGCGCATCACTGCAGACATCTCTCATGTAGTGGTCA 1383

Db 781 GAGACCTCTCTGTCCATGAACCTGGGCCATCTGTGGCCGACATTTCTGTACGTGTGTGATC 840
Qy 1384 CCCACGGGGCGCGCACTGCGGTGGCTTGAGAGCTTCACTCCACCTGCTGGGGGAC 1443
Db 841 CCTACCGAGCGTCCACCGCGAGGCTTCCAGATCGTGTCTCCACCTGCTGGGTGAT 900
Qy 1444 GCGGGAGCCCTACCTCATTTGGCTTTATCTCAGACCTGATCCGCCAGAGCACTAAGGAC 1503
Db 901 GCTGGAGCCCTACCTCATTTGGCTGATCTCTACCGCCCTGCGCGGAACTGGGCCCCC 960
Qy 1504 TCCCGGCTGCGGAGTTCTTGGGCTGAGGCTGGGCTACGGGCTCATGCTTGGCCCTTGTCTGTG 1563
Db 961 TCCTTCTTCTGGAGTTTCCGGGCTCTGAGTCTCTCGCTCATGCTCTGCGCGTTGTGTGG 1020
Qy 1564 GTCTGGGGCGCATGTTCTCTCCGCACTGCGGCTCTTCTTCTGTCAGGACCGCGCCAGG 1623
Db 1021 GCATGGGGCG--GCATTTCTCTGGGCAACCGCACTTCTATTGAGCGCGACCGCGGG 1078
Qy 1624 GCTGAGCAGCACTGGGGG 1642
Db 1079 GCACAGCTGCAGTGCAGG 1097

RESULT 4

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ59pt-fls
; US-08-232-463-14

Query Match 3.4%; Score 69.4; DB 1; Length 7218;

[illegible]

RESULT 5
US-09-616-289-48/c
; Sequence 48, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616.289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 2561
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (246)...(1895)
US-09-616-289-48

```

Qy  204  GCTGTGGGAGCGTAAAGCGGGCCCGGACCGGACCCCGCCCGGCGAC 257
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  587  GCGCGGCGGGGCGGCGGGGTGGCTCGCGCGCGGGCGGCTGGACGCGGCGCG 534

RESULT 6
US-09-128-155-16/c
; Sequence 16, Application US/09128155
; Patent NO. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; TITLE OF INVENTION, AND USES THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

```

RESULT 7
US-09-252-991A-15609
; Sequence 15609, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 15609

; LENGTH: 1560

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-15609

Query Match 2.9%; Score 60.2; DB 4; Length 1560;

Best Local Similarity 47.9%; Pred. No. 0.0024;

Matches 241; Conservative 0; Mismatches 253; Indels 9; Gaps 2;

Qy 260 CCGGCTGGCAGCTACTGCAAGGCGCCCGCGCTCAGCAGCCCAAAACCGGCAGCTTGG 319

Db 3 CCGGCTGGCTTGGGTTCCGTCGCCCGCCAGGTGCGCTTGTCCGATCTCGCCAGAAC 62

Qy 320 GCCCGGGCGGGGCGGCGCCGCCATCTTCCAGCTTGGGCAAGTGTCTCAACTACCTGG 379

Db 63 CGCCCGCGCTCGGTACCGTCTCTCGCGGCTGATGCTCGATCTTCTCAGCGCCCTCG 122

Qy 380 ACAGGTACACGCTGGCAGGCTCTCTTGGACATCCAGACGACTTTGGGGTCAAGAAC 439

Db 123 ACCAGACCATGCTGCGGCTCTCCCTGCGGCGATCTCGGCGACTTCGCGACCTCGACC 182

Qy 440 GAGCGCGGCGCTCTCAGCTCAGTGTTCATCTAGCTTTCATGCTGGCTGCCCATCT 499

Db 183 TGTGCGCTG---GGTGATCTCGGCTTACATGTTGGCGATGACCGTATCGATGCCGATCT 239

Qy 500 TCGGCTACCTTGGGCGACCGCTTCAACAGGAAGTGTCTCAGCTGCGGCAATTTCTTCT 559

Db 240 ACGGCAAGCTCGGGGACTCTTACGGCGCGCGCTGTGATGCTCTTCGCCATCTCGTGT 299

Qy 560 GGTGCGCGCTACCTTCTCCAGCTCTTCAATCCCGACGAGTACTTCTGGCTGTGTGCC 619

Db 300 TCACCGCGGCTCGCTCTTGGGGCTCGC-----GCAGAGCATGGGCCAACTGGTGC 353

Qy 620 TGTCCCGGGGCTGTGGGATCGGGGAGGCGGAGGCGGCTACTCCACATCGCCCCACTATCA 679

Db 354 TGGCGCGAGTGTGTCAGGGGATCGCGCGGCGGCTGATGGCGGTGAGCCAGGCGATCA 413

Qy 680 TTGGCGACCTTTCACCAAGAACACGCTAGCTGCTGCTCGTCTTCTTACTTCGCCA 739

Db 414 TCGGCGACATCGTCCCGCGGCGAGCGCGGCGCTTACCGGGTTACTTCAGCAGCATGT 473

Qy 740 TCCCACTGGGCGAGTGGCTGGGC 762

Db 474 ACGCATCGCAGGCTCGCCGCG 496

RESULT 8

US-09-252-991A-15645

; Sequence 15645, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 15645

; LENGTH: 1680

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-15645

Query Match

Best Local Similarity 2.9%; Score 60.2; DB 4; Length 1680;

Matches 241; Conservative 0; Mismatches 253; Indels 9; Gaps 2;

Qy 260 CCGGCTGGCAGCTACTGCAAGGCGCCCGCGCTCAGCAGCCCAAAACCGGCAGCTTGG 319

Db 1877 CCGGCTGGCTTGGGTTCCGTCGCCCGCCAGGTGCGCTTGTCCGATCTCGCCAGAAC 1818

Qy 320 GCCCGGGCGGGGCGGCGCGCCCATCTCAGCTTGGGCAAGTGTCTCAACTACCTGG 379

Db 1817 CGCCCGCGCTCGGTACCGTCTCGCGGCGCTGATGCTGTCGATCTTCTCCTCAGGCGCTCG 1758

Qy 380 ACAGGTACACCGTGGGCGGCTCTTCTGACATCCAGCAGCATTTTGGGGTCAAGGACC 439

Qy 260 CCGGCTGGCAGCTACTGCAAGGCGCCCGCGCTCAGCAGCCCAAAACCGGCAGCTTGG 319

Db 137 CCGGCTGGCTTGGGTTCCGTCGCCCGCCAGGTGCGCTTGTCCGATCTCGCCAGAAC 196

Qy 320 GCCCGGGCGGGGCGGCGCCGCCATCTCAGCTTGGGCAAGTGTCTCAACTACCTGG 379

Db 197 CGCCCGCGCTCGGTACCGTCTCTCGCGGCGCTGATGCTGTGATCTTCTCAGCGGCTCG 256

Qy 380 ACAGGTACACCGTGGCAGGCGTCTCTTGGACATCCAGACGACTTTGGGGTCAAGGACC 439

Db 257 ACCAGACCATGTCGCCGCTTCCCTGCGGCGATCTCGGCGACTTCGCCGACCTCGACC 316

Qy 440 GAGCGCGGCGCTGCTGCGAGTCACTGTTCATCTAGCTTTCATGCTGGCTGCCCATCT 499

Db 317 TGTGCGCTG---GGTGATCTCGGCTTACATGTTGGCGATGACCGTATCGATGCCGATCT 373

Qy 500 TCGGCTACCTTGGGCGACCGCTTCAACAGGAAGTGTCTCAGCTGCGGCAATTTCTTCT 559

Db 374 ACGGCAAGCTCGGGGACTCTTACGGCGCGCGCTGTGATGCTCTTCGCCATCTCGCTGT 433

Qy 560 GGTGCGCGCTACCTTCTCCAGCTCTTCAATCCCGACGAGTACTTCTGGCTGTGTGCC 619

Db 434 TCACCGCGGCTCGCTCTTGGGGCTCGC-----GCAGAGCATGGGCCAACTGGTGC 487

Qy 620 TGTCCCGGGGCTGTGGGATCGGGGAGGCGGAGGCTACTCCACATCGCCCCACTATCA 679

Db 488 TGGCGCGAGTGTGCGAGGATCGGCGGCGGCGGCTGATGGCGGTGAGCCAGGCGATCA 547

Qy 680 TTGGCGACCTTTCACCAAGAACACGCTAGCTGCTGCTGCTTCTTCTTACTTCGCCA 739

Db 548 TCGGCGACATCGTCCCGCGGCGAGCGCGGCGCTACCGGGTTACTTCAGCAGCATGT 607

Qy 740 TCCCACTGGGCGAGTGGCTGGGC 762

Db 608 ACGCATCGCAGGCTCGCCGCG 630

RESULT 9

US-09-252-991A-15709/c

; Sequence 15709, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 15709

; LENGTH: 1899

; TYPE: DNA

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-15709

Query Match

Best Local Similarity 2.9%; Score 60.2; DB 4; Length 1899;

Matches 241; Conservative 0; Mismatches 253; Indels 9; Gaps 2;

Qy 260 CCGGCTGGCAGCTACTGCAAGGCGCCCGCGCTCAGCAGCCCAAAACCGGCAGCTTGG 319

Db 1877 CCGGCTGGCTTGGGTTCCGTCGCCCGCCAGGTGCGCTTGTCCGATCTCGCCAGAAC 1818

Qy 320 GCCCGGGCGGGGCGGCGCGCCCATCTCAGCTTGGGCAAGTGTCTCAACTACCTGG 379

Db 1817 CGCCCGCGCTCGGTACCGTCTCGCGGCGCTGATGCTGTCGATCTTCTCCTCAGGCGCTCG 1758

Qy 380 ACAGGTACACCGTGGGCGGCTCTTCTGACATCCAGCAGCATTTTGGGGTCAAGGACC 439

Qy 231 GACCGGACCCCGCGGACCCCG 253
Db 107 GGNNTCTCCAGCGCGCGCTCG 85

RESULT 13

US-09-616-289-45/c
; Sequence 45, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 1614
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1614)
US-09-616-289-45

Query Match 2.9%; Score 59.8; DB 4; Length 1614;
Best Local Similarity 52.2%; Pred. No. 0.0029;
Matches 133; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

Qy 6 CGCGCGCCCGATCCGCGCGCGCATGATGCTCGAATGCGCTCGCGCGCGCGCGCGG 65
Db 540 CGCGCGCGCCAGGGCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCA 481
Qy 66 CGCGGAGGAGGAGCGGACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
Db 480 GGGGGGCGCTGTGC 421
Qy 126 GGCTGGCGGTAGCGTTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
Db 420 GACGGGCGCGCGCGGTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 361
Qy 186 GGGCGATGAGTGCAGACGCTGTGCGGCGAGCGTAAAGCGGCGCGCGCGCGCGCG 245
Db 360 GCGGGGCG 301
Qy 246 CACCCCCCGGACCCCC 260
Db 300 CTGACGCGCGCGCGC 286

RESULT 14

US-09-616-289-50/c
; Sequence 50, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.

; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 12425
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-50

Query Match 2.9%; Score 59.8; DB 4; Length 12425;
Best Local Similarity 52.2%; Pred. No. 0.0045;
Matches 133; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
Qy 6 CGCGCGCCCGATCCGCGCGCGCATGATGCTCGAATGCGCTCGCGCGCGCGCGCGG 65
Db 3371 CGCGCGCGCCAGGGCGCGCGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCA 3312
Qy 66 CGCGGAGGAGGAGCGGACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 125
Db 3311 GGGGGGCGCTGTGC 3252
Qy 126 GGCTGGCGGTAGCGTTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 185
Db 3251 GACGGGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3192
Qy 186 GGGCGATGAGTGCAGACGCTGTGCGGCGAGCGTAAAGCGGCGCGCGCGCGCGCG 245
Db 3191 GGGGGGCG 3132
Qy 246 CACCCCCCGGACCCCC 260
Db 3131 CTGACGCGCGCGCGC 3117

RESULT 15

US-09-976-594-1091
; Sequence 1091, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchsinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 1091
; LENGTH: 2030
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 221042.1
US-09-976-594-1091

Query Match 2.9%; Score 59.2; DB 4; Length 2030;
Best Local Similarity 49.1%; Pred. No. 0.004;

| Result No. | Score | Query | | DB | ID | Description |
|---------------|-------|-------|--------|----|--------------------|--------------------|
| | | Match | Length | | | |
| 1 | 2071 | 100.0 | 2071 | 16 | US-10-085-198-47 | Sequence 47, Appl |
| 2 | 1525 | 73.7 | 3179 | 15 | US-10-168-651-36 | Sequence 36, Appl |
| 3 | 563.8 | 27.2 | 1587 | 13 | US-10-024-623-6 | Sequence 6, Appl |
| 4 | 563.8 | 27.2 | 1587 | 15 | US-10-154-419-56 | Sequence 56, Appl |
| 5 | 563.8 | 27.2 | 1587 | 15 | US-10-146-733-51 | Sequence 51, Appl |
| 6 | 563.8 | 27.2 | 2165 | 17 | US-10-648-593-127 | Sequence 127, Appl |
| 7 | 563.8 | 27.2 | 2218 | 13 | US-10-098-841-56 | Sequence 56, Appl |
| 8 | 563.8 | 27.2 | 2230 | 13 | US-10-024-623-4 | Sequence 4, Appl |
| 9 | 563.8 | 27.2 | 2230 | 15 | US-10-154-419-54 | Sequence 54, Appl |
| 10 | 563.8 | 27.2 | 2230 | 15 | US-10-146-733-49 | Sequence 49, Appl |
| 11 | 563.8 | 27.2 | 2264 | 9 | US-09-833-381-883 | Sequence 883, App |
| 12 | 562.6 | 27.2 | 2166 | 16 | US-10-364-231-1236 | Sequence 1236, App |

! TYPE: DNA
! ORGANISM: Homo sapiens
US-10-085-198-47

Query Match 100.0%; Score 2071; DB 16; Length 2071;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2071; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCCCCGCCCGCATCCGGCCCGGCATGATGCTGGAATGCGCGCTCGGGCGGGG 60
DB 1 CCCCCCGCCCGCATCCGGCCCGGCATGATGCTGGAATGCGCGCTCGGGCGGGG 60
QY 61 GGGCGCGGAGGAGAGAGCGGACGCGGAGCGGCGCGCGCGCGCGCGGGGCGCAG 120
DB 61 GGGCGCGGAGGAGAGAGCGGACGCGGAGCGGCGCGCGCGCGCGGGGCGCAG 120
QY 121 CGAGGGGCTGGCGGTAGCGGTTGCTGCGGGGCGCGGGGCGCGCGCTGAGGTTCTG 180
DB 121 CGAGGGGCTGGCGGTAGCGGTTGCTGCGGGGCGCGGGGCGCGCGCTGAGGTTCTG 180
QY 181 GCCCGGGGATGAGGTGCGAGCGCTGCGGCGAGCGTAAGGGGGCGCGGACCGACCC 240
DB 181 GCCCGGGGATGAGGTGCGAGCGCTGCGGCGAGCGTAAGGGGGCGCGGACCGACCC 240
QY 241 CCGGGCACCCCGGACCCCGGCTGCGAGCTACTGCAAAAGGGGCGCGGCTCAGCAG 300
DB 241 CCGGGCACCCCGGACCCCGGCTGCGAGCTACTGCAAAAGGGGCGCGGCTCAGCAG 300
QY 301 CCCAAACCGGCGAGCTTGGGCGCGGGGCGGGGCGAGCGCGCCATCTCAGTTGGGC 360
DB 301 CCCAAACCGGCGAGCTTGGGCGCGGGGCGGGGCGAGCGCGCCATCTCAGTTGGGC 360
QY 361 AAGCTGCTCAACTGACGAGGTACCGTGGCAGGCGTCTTCTGACATCCAGCAG 420
DB 361 AAGCTGCTCAACTGACGAGGTACCGTGGCAGGCGTCTTCTGACATCCAGCAG 420
QY 421 CACTTTGGGGTCAAGAACCGAGCGCGGCGCTGCTGCACTAGTGTTCATCTGTAGCTTC 480
DB 421 CACTTTGGGGTCAAGAACCGAGCGCGGCGCTGCTGCACTAGTGTTCATCTGTAGCTTC 480
QY 481 ATGGTGGCTGCCCGCATCTTGGCTACCTTGGCGGACCGGCTTCAACAGGAAGGTGATCTC 540
DB 481 ATGGTGGCTGCCCGCATCTTGGCTACCTTGGCGGACCGGCTTCAACAGGAAGGTGATCTC 540
QY 541 AGCTGGCGCATTTCTTCTGCTGGCGGCTCACTTCTCCAGCTCTCTTCAATCCCGCAG 600
DB 541 AGCTGGCGCATTTCTTCTGCTGGCGGCTCACTTCTCCAGCTCTCTTCAATCCCGCAG 600
QY 601 TACTTCTGGCTGTGCTGCTGCTCCGGGGCTGGTGGGATCCGGGAGGCGGCTACTCC 660
DB 601 TACTTCTGGCTGTGCTGCTGCTCCGGGGCTGGTGGGATCCGGGAGGCGGCTACTCC 660
QY 661 ACCATCGCCCCACTATCACTTGGCGAGCTTTCACCAAGACACGCGTACGCTCATGCTG 720
DB 661 ACCATCGCCCCACTATCACTTGGCGAGCTTTCACCAAGACACGCGTACGCTCATGCTG 720
QY 721 TCCGTTCTTCTACTTCCCATCCACTTGGGCACTGGGCTGGGCTAGACTTACTGGCTCCAGC 780
DB 721 TCCGTTCTTCTACTTCCCATCCACTTGGGCACTGGGCTGGGCTAGACTTACTGGCTCCAGC 780
QY 781 GTGAGCAGGACGCGGAGACTGGGCACTGGGGATTTGGGGTGTCCCTGTCTTGGGCATG 840
DB 781 GTGAGCAGGACGCGGAGACTGGGCACTGGGGATTTGGGGTGTCCCTGTCTTGGGCATG 840
QY 841 ATCACAGGAACACTCATCTCTANTTCTGTTCCCGACCTATAAAGGGGTCAATGCGGACG 900
DB 841 ATCACAGGAACACTCATCTCTANTTCTGTTCCCGACCTATAAAGGGGTCAATGCGGACG 900
QY 901 CTCGGGACAGCTCAAGCGCGGACCTCATGGCTCCGAGATATGAAGGCCCTGATTGGA 960
DB 901 CTCGGGACAGCTCAAGCGCGGACCTCATGGCTCCGAGATATGAAGGCCCTGATTGGA 960
QY 961 AACCGCAGCTAGCTTCTCTCTCTCCCTGGCGCAGCTCGGCTGTCTCTTCGCCACGGGGCC 1020

DB 961 AACCCGAGCTAGCTCTTCTCCTCCCTGGCCACGCTGGCTGTCTCTTCGCCACGGGGCC 1020
QY 1021 CTGGGCATGTGGATCCCGCTCTACCTGCACCGCGCCCAAGTTGTGCAGAGACACAGAG 1080
DB 1021 CTGGGCATGTGGATCCCGCTCTACCTGCACCGCGCCCAAGTTGTGCAGAGACACAGAG 1080
QY 1081 ACGTGCAAACAGCCCGCTGTGGGGCCCAAGGACAGCTCATCTTTGGGGCCATCACTTC 1140
DB 1081 ACGTGCAAACAGCCCGCTGTGGGGCCCAAGGACAGCTCATCTTTGGGGCCATCACTTC 1140
QY 1141 TTTACGGGATTTCTGGGCGTGTACGCGGGGAGAGCCACGCGTGTGGGCTGTGGCTGAG 1200
DB 1141 TTTACGGGATTTCTGGGCGTGTACGCGGGGAGAGCCACGCGTGTGGGCTGTGGCTGAG 1200
QY 1201 ACCCAGCGGGCGGACCCCACTGTGTGCTGGCGATGCTGGGCTCTGCCATCTTCATC 1260
DB 1201 ACCCAGCGGGCGGACCCCACTGTGTGCTGGCGATGCTGGGCTCTGCCATCTTCATC 1260
QY 1261 TGCTGATCTTCTGGCTGCGCAAGACAGCATCTGTAGGAGCTATATCTGTATCTTCGTC 1320
DB 1261 TGCTGATCTTCTGGCTGCGCAAGACAGCATCTGTAGGAGCTATATCTGTATCTTCGTC 1320
QY 1321 GGGGAGACGCTGCTGTTTTCTAACTGGGCCCATCACTGCAGACATCTCATGTAGTGTGTC 1380
DB 1321 GGGGAGACGCTGCTGTTTTCTAACTGGGCCCATCACTGCAGACATCTCATGTAGTGTGTC 1380
QY 1381 ATCCCCACGCGCGCGCACCTGCCGTGGCTTTGCAGAGCTTTCACCTCCACCTGCTGGGG 1440
DB 1381 ATCCCCACGCGCGCGCACCTGCCGTGGCTTTGCAGAGCTTTCACCTCCACCTGCTGGGG 1440
QY 1441 GACCGCGGGAGCCCTTACCTCATTTGGCTTTATCTCAGACCTGATCCGCCAGACACTAAG 1500
DB 1441 GACCGCGGGAGCCCTTACCTCATTTGGCTTTATCTCAGACCTGATCCGCCAGACACTAAG 1500
QY 1501 GACTCCCGCTCTGGGAGTTCTCAGCTGGGCTACCGCTCATGCTCTGCCCTTTGCTC 1560
DB 1501 GACTCCCGCTCTGGGAGTTCTCAGCTGGGCTACCGCTCATGCTCTGCCCTTTGCTC 1560
QY 1561 GTGTCTCTGGCGGCTGTTCTTCTCGCACTGCGCTCTTCTTCGTCAAGCAGCGGCGCC 1620
DB 1561 GTGTCTCTGGCGGCTGTTCTTCTCGCACTGCGCTCTTCTTCGTCAAGCAGCGGCGCC 1620
QY 1621 AGGCTCAGACGACCTTGGGGAGAGACGCGCGGGGTCAAGGTGTGTCATCAGCGGGGG 1680
DB 1621 AGGCTCAGACGACCTTGGGGAGAGACGCGCGGGGTCAAGGTGTGTCATCAGCGGGGG 1680
QY 1681 CCGGGCCCGGGCACTGCTCTGGCACAATCTGTGTGGGGGCCAGCTGACCGGAGGTGCTG 1740
DB 1681 CCGGGCCCGGGCACTGCTCTGGCACAATCTGTGTGGGGGCCAGCTGACCGGAGGTGCTG 1740
QY 1741 GGCAGGAGACTCTGTACCGCCAGGGGAGATGGGAGAGCCAGGGTGGGGAGAGAGAG 1800
DB 1741 GGCAGGAGACTCTGTACCGCCAGGGGAGATGGGAGAGCCAGGGTGGGGAGAGAGAG 1800
QY 1801 AGAGAGAGTAAAGAGAGGAGAGAAAGTCAAGAACTAGAGGAGGAGGGGGGGGGCC 1860
DB 1801 AGAGAGAGTAAAGAGAGGAGAGAAAGTCAAGAACTAGAGGAGGAGGGGGGGGGCC 1860
QY 1861 CAGCTTTGAAAAACACTAAGTCCAGAGACAAACCCAAAGTCTGGATCCACAGACACCCC 1920
DB 1861 CAGCTTTGAAAAACACTAAGTCCAGAGACAAACCCAAAGTCTGGATCCACAGACACCCC 1920
QY 1921 GTGGCCTCCACAGCTCCAGGCTGACCTGGCACTGGGCTCAGGGCTGAGCCCGCAGCAA 1980
DB 1921 GTGGCCTCCACAGCTCCAGGCTGACCTGGCACTGGGCTCAGGGCTGAGCCCGCAGCAA 1980
QY 1981 CCAAGTGGGTGCACTGAGTGCATGGGAGGTCTGTACCTTCCCGCCCGCCACCCAGGGCAGG 2040
DB 1981 CCAAGTGGGTGCACTGAGTGCATGGGAGGTCTGTACCTTCCCGCCCGCCACCCAGGGCAGG 2040
QY 2041 GCTCAGGTGGCTATCAGGTTCCCTGCTTCC 2071

; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 56
; LENGTH: 1587
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1587)
US-10-154-419-56

Query Match 27.2%; Score 563.8; DB 15; Length 1587;

Best Local Similarity 64.3%; Pred. No. 5.6e-133;

Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;

| | | | |
|----|------|---|------|
| QY | 282 | GGGCGCGGCTCAGCAGCCAAACCGCCAGCTTGGCGCGGGGGGGGAGCGCGC | 341 |
| DB | 126 | GGACGAGGGGCTGCAGCGCATCAGCGGCTGTCTCCGGCGTTTCGGCTCTCATAGT | 185 |
| QY | 342 | CGCCATCTCAGCTTGGGCAACGTCTCACTACGACAGGTACACCGTGGCAGCGT | 401 |
| DB | 186 | GGCGGTCTGTCTACATCATCTCTCTGAATACATGACCGCTTACCGTGGCTGGCGT | 245 |
| QY | 402 | CTTCTTGACATCCAGCAGCATTGGGGTCAAGACCGAGCGCGCCCTGTCTGCAGTC | 461 |
| DB | 246 | CTTCCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAG | 305 |
| QY | 462 | AGTGTTCATCTGAGCTTCAATGTGGCTGCCCATCTTCGGTACCTGGGCGACCGCTT | 521 |
| DB | 306 | CGTGTTCATCTCCAGTACATGTGTGGCACCTGTGTGGTACCTGGTGGTACAGGTA | 365 |
| QY | 522 | CAACAGGAAGTGTCTCAGCTGCGCATTTCTCTGTGTCGCGCTCACTTCTCCAG | 581 |
| DB | 366 | CAATCGGAAGTATCTCATGTGCGGGGCAATGCTCTGTGCTCTGTGACACTGGGGT | 425 |
| QY | 582 | CTCCTTTCATCCCGAGCAGTATCTTGGCTGTGCTCTGCTCCCGGGGCTGTGGGCAT | 641 |
| DB | 426 | ATCCTTTCATCCCGAGCAGTATCTTGGCTGTGCTCTGCTCCCGGGGCTGTGGGGT | 485 |
| QY | 642 | CGGGGAGGCGAGTACTCAGCATCGCCCGCAGTATCATTTGGGAGCCTTTCACCAAGAA | 701 |
| DB | 486 | CGGGGAGGCGAGTATTTCCACCATCGCGCCCACTCTCATTTGGGAGCCTTTTGGCGGA | 545 |
| QY | 702 | CAGCGTACGCTCATGTCTGCTCTTCTTACTTCTGCTCCCATCCACTGGGAGTGGCTGG | 761 |
| DB | 546 | CAAGCGAGCGGATGCTCAGCATCTTCTACTTTGCCATTCGGTGGGAGTGGTCTGGG | 605 |
| QY | 762 | CTACATTACTGGCTCCAGCGTGAAGCAGGCGAGACTGGCAGCTGGGCAATGGGGT | 821 |
| DB | 606 | CTACATTGCAAGGCTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGCTCTGAGGT | 665 |
| QY | 822 | GTCCCTGTCTGGGCAATGATCAGAAACATCATCTCTCATTTGTGTCCAGCCACTAA | 881 |
| DB | 666 | GACACCGGCTAGGAGTGGTGGCGTCTGCTGCTGCTTCTGCTAGTGGGAGCGCC | 725 |
| QY | 882 | AAGGGTCTAGCGACAGCTCGGGA---CCAGCTCAAGCCCGGACCTCATGGCTCCG | 938 |
| DB | 726 | AAGGGAGCGTGGAGCGCCACTCAGATTTGGCACCCCTGAAACCCCACTCTGGTGGG | 785 |
| QY | 939 | AGATATGAAGGCGCTGATTCGAAACCGCAGTACGCTTCTCTCTCCCTGGCCACGCTCGG | 998 |
| DB | 786 | AGATCTGAGGCTCTGGCAGAAATCTAGTTTGGTCTGCTCTCTCTGGCTTCTACTGC | 845 |
| QY | 999 | TGCTCTCTTCCGACCGGGGCGCTGGGCAATGTCATCCGCTCTACTGCAACCGCGCCCA | 1058 |
| DB | 846 | TGTGGCTTTGTACCGGGCTCCCTGGCTCTGTGGGCTCCGGCATCTCTGTGGTTCGG | 905 |
| QY | 1059 | AGTTGTGCA---GAAGACAGCAGAGCTGCAACAGCCCGCTGTGGGCGCAAGACAG | 1115 |
| DB | 906 | CGTGGTCTTGGGAGAGCCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 965 |

| | | | |
|----|------|--|------|
| QY | 1116 | CCTCATCTTGGGGCCATCATCTGCTTTACGGGATTTCTGGCGTGTGTACGGGGGAGG | 1175 |
| DB | 966 | TCTCATCTTGGGACTCATCACTGCTGACCGGAGTCTGGGTGTGGGCTGGGTGTGGA | 1025 |
| QY | 1176 | AGCCACGCGTGGTGGCGCTGAAGACCCAGCGGGCCGACCCACTGGTGTGTGGCGTGG | 1235 |
| DB | 1026 | GATCAGCGCGGCTCGGCCACTCAACACCCCGGGCTGATCCCTGGTCTGTGGCAGTGG | 1085 |
| QY | 1236 | CATGCTGGGCTCTGCGCATCTTCTGCTGCTGATCTTCTGCTGCTGCTGCTGCTGCTG | 1295 |
| DB | 1086 | CTTCTGGGCTCTGCGACCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1145 |
| QY | 1296 | AGGAGCCTATATCTGATCTTCTGCTGGGAGAGCTGCTGCTGCTTCTAACTGGGCGCAT | 1355 |
| DB | 1146 | GGCCACTTATATTTTCTATCTTCTGAGAGAGCCCTCTCTGCTGCTGCTGCTGCTGCT | 1205 |
| QY | 1356 | TGCAGACATCTCATGTACGTGCTATCCCAACCGGGGCGGCTGCTGCTGCTGCTGCTG | 1415 |
| DB | 1206 | GGCGGACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1265 |
| QY | 1416 | GAGCTTCACTCCACCTGCTGGGGAGCGCGGGGAGCCCTTCTGCTGCTGCTGCTGCTG | 1475 |
| DB | 1266 | GATGCTGCTGCTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1325 |
| QY | 1476 | AGACCTGATCCGCGAGAGCATAAGGACTCCCGCTCTGGGAGTTCCTGAGCTGGGCTA | 1535 |
| DB | 1326 | TGACGCGCTGCGCGGAACTGGCCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1385 |
| QY | 1536 | CGGCTCATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1595 |
| DB | 1386 | CTGCTCATGCTCTGCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 1445 |
| QY | 1596 | GCTCTTCTCTGCTCAGCGAGCGCGCGGCTGAGCAGCAGCTGGGG 1642 | |
| DB | 1446 | CATCTTCTATGAGGCGGAGCGCGGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTG | 1492 |

RESULT 5

US-10-146-733-51
; Sequence 51, Application US/10146733
; Publication No. US20030165891A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC2327, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: HMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MNI-248
; CURRENT APPLICATION NUMBER: US/10/146,733
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/185,938
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/515,520
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/518,866
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/195,993
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/199,799
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 60/233,537
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/235,018
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/235,059
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18

US-10-648-593-127

| | | | | |
|-----------------------|--------|--|--------|-----------------------------------|
| Query Match | 27.2%; | Score 563.8; | DB 17; | Length 2165; |
| Best Local Similarity | 64.3%; | Prod. No. 5.8e-133; | | |
| Matches | 879; | Conservative | 0; | Mismatches 482; Indels 6; Gaps 2; |
| Qy | 282 | GGGCCCCGGCGCTCAGCAGCGCCAAACCGGCCAGCTTGGGCGCGGGCGGGGGGCGAGCCGC | 341 | |
| Db | 450 | GGACGAGAGGGCTGCAGCGCATCACCGGCTGTCTCCGGCCGTTGGGCTCTCATAGT | 509 | |
| Qy | 342 | CGCCATCCTCAGCTTTGGGCAACGTGCTCAACTACTGGACAGGTACACCGTGGCAGCGGT | 401 | |
| Db | 510 | GGCGGTGCTGTGCTATCATCAATCTCTCTGACTACATGGACCGCTTACCGTGGCTGGCGT | 569 | |
| Qy | 402 | CTTCTCGGCATCCAGCAGCACTTTGGGGTCAAGGACGGAGCCGCCGCTGTGCGATC | 461 | |
| Db | 570 | CTTTCGGCAGATCGAGCAGTCTTTCAACATCGGGCAGTAGTAGCTCTGGGCTCATCCAGAC | 629 | |
| Qy | 462 | AGTGTTCATCTGTAGCTTCATGGTGGCTGCCGCCCATCTTCGGCTACCTGGGCGACCGCTT | 521 | |
| Db | 630 | CGTGTTCATCTCCAGTTACATGGTGTGGCACCTGTGTGGTGTACCTGGGTGACAGGTA | 689 | |
| Qy | 522 | CAACAGGAAGTGATTCAGCTCGCGCATTTTCTCTGTGTGGCCGCTACCTTCTCCAG | 581 | |
| Db | 690 | CAATCGGAAGTATCTCATGTGCGGGGCAATGCCCTCTGTGTCCTGTGTGACACTGGGGTC | 749 | |
| Qy | 582 | CTCCTTCATTTCCCAGCAGTACTTCTGGCTGCTGGTCTGTCTCCCGGGGCTGTGGGCAT | 641 | |
| Db | 750 | ATCCTTCATCCCGGAGAGCATTTCTGGCTGCTCTCTTGACCCGGGGCCTGGTGGGGT | 809 | |
| Qy | 642 | CGGGAGGCAGTACTTCCACATCGGCCCCCACTATCATTTGGCGACCTCTTTCACAAAGAA | 701 | |
| Db | 810 | CGGGGAGCGCAGTTATTTCCACATCGCGGCCCACTCTCATTTGCCACCTCTTTGTGGCCGA | 869 | |
| Qy | 702 | CACGCTPACGCTCACTGTCTCGTCTCTACTTCGCCCATCCCACTGGGCGAGTGGCCTGGG | 761 | |
| Db | 870 | CCAGCGAGCGGATGCTCAGATCTTCTACTTTGCCATTCGGGTGGCAGTGTGCTGGG | 929 | |
| Qy | 762 | CTACATTACTGGGCTCCAGCGGTGAAGCAGGAGCGGAGACTGGCACTGGGCAATTGGCGGGT | 821 | |
| Db | 930 | CTACATTTGCAGGCTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGT | 989 | |
| Qy | 822 | GTCCCTGTCTGGGCATGATTCAGAGAACACTCATCTCATTTGTGTCCAGGCCACTAA | 881 | |
| Db | 990 | GACACCGGGTCTAGGAGTGTGGCCGTTCTGCTGCTTCTGTGTAGTGTGGGAGCGGCC | 1049 | |
| Qy | 882 | AAGGGTCTATGCCGACCACTCGGGGA---CCAGCTCAAGGCCCGGACCTCATGGCTCCG | 938 | |
| Db | 1050 | AAGGGGACGCTGGAGCGCCACTCAGATTTGCCACCCCTGAACCCCACTCTGTGTGGGC | 1109 | |
| Qy | 939 | AGATATGAAGGCCCTGATTCGAAACCCGAGCTAGTCTTCTCTCCCTGGCCACGTCGGC | 998 | |
| Db | 1110 | AGATCTGAGGGCTCTGGCAAGAAATCCTAGTTTGGTCTGTCTTCTCCCTGGGCTTCACTGC | 1169 | |
| Qy | 999 | TGTCTCTCTTCGCCACCGGGGCCCTGGGCATGTGATCCCGCTCTACTCTGCAACCGCGCCA | 1058 | |
| Db | 1170 | TGTGGCCTTTGTCACGGGCTCCCTGTCTGTGGGCTCTCGGCATTCCTCTGTGGTTCGCG | 1229 | |
| Qy | 1059 | AGTTGTGCA---GAGACACAGAGAGCTGCAACAGCCCCCCTGTGGGGCCCAAGGACAG | 1115 | |
| Db | 1230 | CGTGGTCTTTGGGGAGACCCACCCCTGCCCTTCCCGGAGACTCTCTGTCTTCTCTGCAG | 1289 | |
| Qy | 1116 | CCTCATCTTTGGGGCCATCACTGTCTTACGGGAATTTCTGGGCTGTGTACAGGGGCGAG | 1175 | |
| Db | 1290 | TCTCATCTTTGGACTCATCACTTGCCTGACCGGAGTCTCTGGGTGTGGGCTTGGGTGTGGA | 1349 | |
| Qy | 1176 | AGCCACGCGCTGTGGCGCTCTGAAGACCCAGCGGGCCGACCCCACTGGTGTGTGCCGTGGG | 1235 | |
| Db | 1350 | GATCAGCGCGCGCTCCGCCACTCCAAACCCCGGGCTGATCCCTGGTCTGTGCCACTGG | 1409 | |
| Qy | 1236 | CATGCTGGGCTCTGCCATCTTCAATCTGCCTGTATCTTCGTGGGCTGCCAAGACAGCATCGT | 1295 | |
| Db | 1410 | CTCTCTGGGCTCTGCACCTCTCTCTTCTGTCTCTTGTCTCTTGTCTGGCCCGCTGTAGCATCGT | 1469 | |

RESULT 7

```

RESULTS
; Sequence 56, Application US/10098841
; Publication No. US20020197679A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Xu, Chongjun
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhang, Jie
; APPLICANT: Qian, Xisichong B.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2
; CURRENT APPLICATION NUMBER: US/10/098,841
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 09/598,042
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 331
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 56
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (372)..(1958)
US-10-098-841-56

```

Query Match 27.2%; Score 563.8; DB 13; Length 2218;
Best Local Similarity 64.3%; Pred. No. 5.8e-133;
Matches 879; Conservative 0; Mismatches 482; Indels 6;

```
QY 282 GGGCCCCGGCGCTCAGCAGCCCAAAACCGGCGAGCTTGGGCGCGGGGGGCGAGCCGC 341
Db 497 GGACAGGAGGGGCTGCGACGCGCATACCGGCGCTGTCTCCCGGCGCTCGGCTCTCATAGT 556
QY 342 CGCCATCTCAGCTTGGGCAACGCTCTCACTACCTGACAGGTACACCGTGGCAGGGGT 401
Db 557 GCGGCTGTCTGCTACATCAATCTCTGAACTACATGAGACCGCTTCAACGCTGGCTGGGT 616
QY 402 CTTTGGACATCCAGCAGCACTTTGGGCTCAAGGACCGAGGCGCGGCTGCTCAGTC 461
Db 617 CTTTCCGACATCGAGCAGTTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCCAG 676
QY 462 AGTGTTCATCTAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 521
Db 677 CGTGTTCATCTCAGTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 736
QY 522 CAACAGGAAGGTGATTCTCAGCTGGGCAATTTTCTTCTGCTGGCGCTCACCTTCTCCAG 581
Db 737 CAATCGGAAGTATCTCATGTGGGGGCAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 796
QY 582 CTCCTTCATTTCCCGCAGCTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
Db 797 ATCTTTCATCCCGGAGAGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 856
QY 642 CGGGGAGGCGAGTACTTCCACCATCGCCCCCACTATCATTTGGGACCTCTTCCACCAAG 701
Db 857 CGGGGAGGCGAGTATTTCACCATCGCCCCCACTCTCATTTGGGACCTCTTTCGCGCA 916
QY 702 CAGCGCTAGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 761
Db 917 CAGCGGAGCGGATGCTCAGCATCTTCTACTTGGCATCCCACTGGGCACTGGGCTGG 976
QY 762 CTACATTTACTGCTCAGCGTGAAGCAGGAGCCGAGAGTGGCACTGGGCACTGGGCT 821
Db 977 CTACATTTGAGGCTCCAAAGTGAAGATATGCTGAGAGCTGGCACTGGGCTCTGAGG 1036
QY 822 GTCCCTGTCTGGGATGATCACAGGAACATCATCTCTCTCTCTCTCTCTCTCTCTCTCT 881
Db 1037 GACACCGGGCTAGGAGTGGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096
QY 882 AAGGGGTATGCGCAGCTCGGGA---CCAGCTCAAGGCGCGGACCTCATGGCTCG 938
Db 1097 AAGGGGAGCGCTGAGGGCCACTCAGATTGCGCACCCCTGAAACCCACCTGCTGGTGG 1156
QY 939 AGATATGAAGGCGCTGATTCGAAACCGCAGCTAGTCTTCTCTCTCTCTCTCTCTCTCTCT 998
Db 1157 AGATCTGAGGCTCTGGCAAGAAATCTAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216
QY 999 TGTCTCTCTCGCACGGGGCGCTGGGCACTGGATCGGATCCCGCTCTACTGCAACCGGCCCA 1058
Db 1217 TGTGGCCTTTGTCACGGGCTCCCTGGCTCTGTGGGCTCGGCACTCTGCTGGGTTCCCG 1276
QY 1059 AGTTGTGCA---GAAGACAGCAGAGAGCTGSCAACAGCGCGCCCTGTGGGCGCAAGACAG 1115
Db 1277 CGTGGTCTTTGGGAGAGCCCAACCTCGCTTCCCGGAGACTCTCTGCTCTCTCTCTGACAG 1336
QY 1116 CTTCACTTTGGGGCCATCATCTGCTTTACGGGATTTCTGGGCGTGGTCAAGGGGCGAG 1175
Db 1337 TCTCATCTTTGGACTCATCACTCTGCTGAGCGGAGTCTGGGTGTGGGCGCTGGGTGGA 1396
QY 1176 AGCACGCGTGGTGGCGCTGAAGACCCAGCGGGCGACCCACTGGTGTGCTGGTGG 1235
Db 1397 GATCAGCGCGGCTCGGCCCATCTCAACCCCGGGGCTGATCCCTGCTGTGCTGCTGCTGCT 1456
QY 1236 CATGCTGGGCTCTGCCATCTTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1295
Db 1457 CTTCTGGGCTCTGACACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1516
QY 1296 AGGAGCCTATCTGTATCTTCTGCTGGGAGACCTGCTCTCTTCTTCACTGGGCCATCAC 1355
Db 1517 GGCACCTTATATTTTCTATCTTCTATTTGAGAGACCTCTCTGTCATGAACTGGGGCATCT 1576
QY 1356 TGCAGACATCTCTATGATGCTGCTGCTATCCCAACCGCGGCGGCCACTGCCGCTGGCTTGA 1415
```

```
Db 1577 GGCCGCAATTCGTGTAAGTGGTATCCCTACCGAGCTTCCACCGCGAGGCGCTTCCA 1636
QY 1416 GAGCTTCACTCCCACTGCTGGGAGAGCGCGGAGAGCCCTACCTGATTTGCTTTATCTC 1475
Db 1637 GATCGTGTGCTCCCACTGCTGGTGTATGCTGGGAGCCCTTACCTCATTTGGCCTGATCTC 1696
QY 1476 AGACTGATCGCCAGAGCACTAAGGACTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTA 1535
Db 1697 TGACCGCTGCGCCGGAACCTGGCCCCCTCTCTTCTTCCGAGTTCGGGCTCTGCAGTT 1756
QY 1536 CGCGCTCATGCTGCTGCCCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1595
Db 1757 CTGCTCATGCTCTGCGGTTTGTGGGCACTGGGCGCGCAGCTTCTCTGGGCAACGCG 1816
QY 1596 GCTCTTCTGCTAGCAGCAGCGCGCAGGCTGAGCAGCAGCCTTGGGG 1642
Db 1817 CATCTTATTGAGCCGACCGCGCGGGCACAGCTGCACGTGAGG 1863

RESULT 8
US-10-024-623-4
; Sequence 4, Application US/10024623
; Publication No. US20020187524A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Roy A. J.
; TITLE OF INVENTION: 8099, 46455, 54414, 53763, 67076, 67102, 44181,
; TITLE OF INVENTION: 67084FL, AND 67084 ALT, HUMAN PROTEINS AND METHODS OF
; FILE REFERENCE: USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/024,623
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2230
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (376)...(1962)
US-10-024-623-4

Query Match 27.2%; Score 563.8; DB 13; Length 2230;
Best Local Similarity 64.3%; Pred. No. 5.8e-133;
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;
```


| | | | | | |
|--|---|--|------|---|------|
| D | b | | 1101 | AAGGGAGCGTGGAGGCGCACCTCAGATTTCGCACCCCTGAACCCCACTCTGTGTGGGC | 1160 |
| Q | y | | 939 | AGATATGAAGGCCCTGATTGAAACCAGCAGCTACGTCTTCTCTCCCTGGCCACGCTCGGC | 998 |
| D | b | | 1161 | AGATCTGAGGGCTCTGGCAAGAATCCTAGTTTCGTCTCTCTCCCTGGGCTTCACTGC | 1220 |
| Q | y | | 999 | TGTCCTCTTGCGCACGGGGGCCCTGGGCATGTGGATCCCGCTCTACTGCACGGGCCCA | 1058 |
| D | b | | 1221 | TGTGGCCTTTGTCAACGGGCTCCCTGGCTCTGTGGGCTCGGGCATTCCTGTGCGTTCGG | 1280 |
| Q | y | | 1059 | AGTTGTGCA--GAAGACGACAGAGCGTGCAACGCCCGCCCTGTGGGCGCAAGACAG | 1115 |
| D | b | | 1281 | CGTGGCTCTGGGGAGACCCACCTTCCTCCGGAGACTCTGTCTCTCTCTGACAG | 1340 |
| Q | y | | 1116 | CCTCATCTTTGGGGCCATCACTGCTTTTAGGGATTTCTGGGGCGTGTCA CGGGGCGAGG | 1175 |
| D | b | | 1341 | TCTCATCTTTGGACTCATCACTTCCTGACCGGAGTCTTGGGTGTGGGCTGGGTGTGGA | 1400 |
| Q | y | | 1176 | AGCAGCGCTGTGGCGCTGAAGACCCAGCGGGCGGACCACTGCTGTGTGCGGTGG | 1235 |
| D | b | | 1401 | GATCAGCGCGCGCTCGCCCATCTCAAACCCCCGGGCTGATCCCCTGTCTGTGCCACTG | 1460 |
| Q | y | | 1236 | CATGCTGGGCTCTGCCATCTTTCATCTCTGCTGATCTTCTGGGCTGCCAAGACGACATCGT | 1295 |
| D | b | | 1461 | CCTCCTGGGCTCTGCACCCCTCTCTCTCTCTCCCTTGCCCTGCGCCGCTGATCGT | 1520 |
| Q | y | | 1296 | AGAGCCTATATCTGTATCTTCTCGGGAGAGCGTCTGTTTTTAATGGGGCATCAC | 1355 |
| D | b | | 1521 | GGCCACTTATATTTTTCATCTTCAATGGAGAGACCTCTCTGTCATGAACCTGGGCCATCGT | 1580 |
| Q | y | | 1356 | TGCAGACATCCTCATGTACGTGTGTCATCCCGCGCGCGCCACCTGCCGTGGCCTTGCA | 1415 |
| D | b | | 1581 | GGCCGACNTTCTGCTGTACGTGTATCCCTAACCCGACGCTCCACCGCGAGSCCTTCCA | 1640 |
| Q | y | | 1416 | GAGCTTCACCTCCACCTCTGTGGGGAGCGCGGGAGCCCTACTCATTTGGCTTTATCTC | 1475 |
| D | b | | 1641 | GATCGTCTCTGCCACCTGCTGGGTGATGTGGGAGCCCTTACCTCATTTGGCCTGATCTC | 1700 |
| Q | y | | 1476 | AGACTCATCGCCAGACGACATAAGGACTCCCGCTCTGGGAGTTCTGAGCTGGGCTA | 1535 |
| D | b | | 1701 | TGACCGGCTGGCGCGAACTTGGCCCCCTCTCTTGTCTCGAGTTCGGGCTCTGCAGTT | 1760 |
| Q | y | | 1536 | CGGCTCATGCTCTGCCCTTTTCGTCTGTGTCCTGGGCGGCACTTCTCTCGCCACTGC | 1595 |
| D | b | | 1761 | CTGCTCATGCTCTGCGCTTTGTGGGCACTGGGCGGCGAGCCCTTCTGGGCAACCGC | 1820 |
| Q | y | | 1596 | CTCTTCTTGTGACGACCGCGCCAGGGCTGAGCAGCACCTGGGG | 1642 |
| D | b | | 1821 | CATCTTCATTGAGGCCGACCGCGCGGGGACAGCTGCAGTGAGG | 1867 |
| RESULT 10 | | | | | |
| US-10-146-733-49 | | | | | |
| ; Sequence 49, Application US/10146733 | | | | | |
| ; Publication No. US20030165891A1 | | | | | |
| ; GENERAL INFORMATION: | | | | | |
| ; APPLICANT: Glucksmann, Maria A. | | | | | |
| ; APPLICANT: Silos-Santiago, Inmaculada | | | | | |
| ; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23327, TWIK-8, IC47615, | | | | | |
| ; TITLE OF INVENTION: HNMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES | | | | | |
| ; TITLE OF INVENTION: AND USE THEREOF | | | | | |
| ; FILE REFERENCE: MNI-248 | | | | | |
| ; CURRENT APPLICATION NUMBER: US/10/146,733 | | | | | |
| ; CURRENT FILING DATE: 2002-05-15 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/185,938 | | | | | |
| ; PRIOR FILING DATE: 2000-02-29 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 09/515,520 | | | | | |
| ; PRIOR FILING DATE: 2000-02-29 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 09/518,866 | | | | | |
| ; PRIOR FILING DATE: 2000-03-03 | | | | | |
| ; PRIOR APPLICATION NUMBER: US 60/195,734 | | | | | |
| ; PRIOR FILING DATE: 2000-04-07 | | | | | |

Db 861 CGGGAGGCCAGTATTATCCACCATCGCGCCCACTCTCATTTGCCGACCTCTTTTGTGGCGGA 920
Qy 702 CACGGTACGCTCATGCTGTCGGTCTTACTTTCGCCATCCCACTGGGCACTGGCCTGGG 761
Db 921 CCAGCGAGCGGATGCTCAGCATCTTCTACTTTGGCAATCCGGTGGGCACTGGTCTGGG 980
Qy 762 CTACATTTACTGGCTCCAGCGTGAAGCAGGAGCGGAGACTGGGCACTGGGCACTGGGCT 821
Db 981 CTACATTTGAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGT 1040
Qy 822 GTCCCTCTCTCGGCACTGATCACAGGAACAATCATCTCATTTCTGGTCCCGACCACTAA 881
Db 1041 GACACCGGCTTAGAGTGGTGGCGTCTCTGCTGCTTCTGCTAGTGGCGGAGCGGCC 1100
Qy 882 AAGGGGTCAACCGACCACTGCGGGA---CCAGCTCAAGGCCCGGACCTCATGGCTCCG 938
Db 1101 AAGGGAGCGGTGGAGGCCACTCAGATTTGCCACCCCTGAACCCCACTCTGCTGGGC 1160
Qy 939 AGATATGAGGCCCTGATTGCAAAACCGAGCTAGCTTCTCTCTCTCTCTCTCTCTCTCT 998
Db 1161 AGATCTGAGGGCTCTGGCAAGAAATCTTAGTTTCTGCTGCTCTCTCTCTCTCTCTCT 1220
Qy 999 TGTCTCTCTCGCAGCGGGGCTGGGCACTGATGATCCGCTCTCTCTCTCTCTCTCTCTCT 1058
Db 1221 TGTGGCTTTGTCAAGGCTCTCTGCTCTCTGCTCTCTGCTCTCTCTCTCTCTCTCTCT 1280
Qy 1059 AGTTGTGCA---GAAGCAGCAGAGAGCGTGCAACAGCGCGCTCTGGGGCAAGGACAG 1115
Db 1281 CGTGTCTTTGGGAGAGCCCACTGCTTCCCGAGACTCTCTCTCTCTCTCTCTCTCTCT 1340
Qy 1116 CCTCATCTTTGGGGCACTACCTGCTTTAAGGATTTCTGGGCGTGTGTCACGGGGGAGG 1175
Db 1341 TCTCATCTTTGAGACTCATCTGCTGACCGGAGTCTGGGCTGGGCGCTGGGCTGTGGA 1400
Qy 1176 AGCCACGCTGTGCGCGCTGAAGACCCAGCGGCGGACCACTGCTGTGTGCTGGG 1235
Db 1401 GATCAGCGCGCGCTCGCGCACTCAACCCCGGGCTGATCCCTGCTGTGTGCTGCTG 1460
Qy 1236 CATCTGGGCTCTGCCATCTTCACTGCTGCTGATCTTGTGCTGCTGCTGCTGCTGCTG 1295
Db 1461 CCTCTGGGCTCTGACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1520
Qy 1296 AGGAGCTATATCTGTATCTTGTGCTGGGAGACGCTGCTGCTTCTAACTGGGCACTAC 1355
Db 1521 GGCCACTTATATTTTCACTTCACTGAGAGACCTCTCTGCTCAATGAATGGGCACTGT 1580
Qy 1356 TGCAGACATCTCATGTAGTGTGATCCCAACCGCGGCGGCACTGCTGCTGGCTTGCA 1415
Db 1581 GGCCGACATTTGCTGTACGCTGTGATCCCTACCGAGCTTCCACCGCGGCGCTTCCA 1640
Qy 1416 GAGCTTCACTCCACCTGCTGGGGAGCGCGGAGCGCCCTTACTCATTTGCTTTATCTC 1475
Db 1641 GATGCTGCTGCTCCACCTGCTGGGTGATGCTGGGAGCGCCCTTACTCATTTGCTGATCTC 1700
Qy 1476 AGACTGTATCCGCGAGACATAAGGACTCCCGCTCTGGGAGTTCCTGAGCCTGGGCTA 1535
Db 1701 TGACCGCTTGCGCGGAACCTGGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1760
Qy 1536 CGGCTCATGCTCTGCGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1595
Db 1761 CTGCTCATGCTCTGCGCTTTGTTGGGCACTGGGCGGCGAGCCCTTCTCTGCGCACCGC 1820
Qy 1596 GCTCTCTCTGTCAGACCGCGGCTGAGGCTGAGCAGCACTGGGG 1642
Db 1821 CATCTTCAATTGAGCGCCAGCGCGGCGGAGCAGCTGACGTGAGG 1867

RESULT 11

US-09-833-381-883

; Sequence 883, Application US/09833381

; Patent No. US20020132090A1

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 883
; LENGTH: 2264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2264)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-883

Query Match 27.2%; Score 563.8; DB 9; Length 2264;
Best Local Similarity 64.3%; Pred. No. 5.8e-133;
Matches 879; Conservative 0; Mismatches 482; Indels 6; Gaps 2;
Qy 282 GGGCCCCGGCGCTCAGCAGCCCAACCGGCCAGCTTGGGCGCGGGGGGGGAGCGCCG 341
Db 516 GAGCCAGGAGGGGCTGCAGCGCATCACCGGCTGTCTCCGCGCGTTCCGCTCTCATAGT 575
Qy 342 CGCCATCTCAGCTTGGGCAACGTGCTCACTCTGACAGGTACACCGTGGGAGGGGT 401
Db 576 GGGGCTCTGTCTATCATATCTCTGAACTATACATGACCGCTTACCGTGGCTGGCGT 635
Qy 402 CCTTCTGGACATCCAGCAGCACTTTGGGGTCAAGGACCGAGGCGCGCTGTGCTGCAATC 461
Db 636 CCTTCCCGACATCGAGCAGTTCTTCAACATCGGGGACAGTAGTCTTGGGCTCATCCAGAC 695
Qy 462 AGTGTTCATCTGAGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 521
Db 696 CGTGTTCATCTCAGTTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
Qy 522 CAACAGGAAGCTGATCTCAGCTCGGCACTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581
Db 756 CATTCGGAAGTATCTCATGTGCGGGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
Qy 582 CTCCTTTCATTTCCAGCAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
Db 816 ATCTTTCATCTCCCGAGAGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 875
Qy 642 CGGGAGGCGAGCTACTTCCACATGCGCCCACTATCATTTGCGGAGCTCTTTCACAAAGAA 701
Db 876 CGGGAGGCGAGTTATTCCACCATCGCGCCACTCTCATTTGCGGAGCTCTTTCGCGCGCA 935
Qy 702 CAGCGCTACGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 761
Db 936 CCAGCGAGCGGATGCTCAGCATCTTCTACTTTGCCATTTCCGTTGGGAGTGGTCTGCGG 995
Qy 762 CTACATTTACTTGGCTTCAGCGCTGAAGCAGGCGAGAGCTGCGACATTTGGGCACTTCGCGGT 821
Db 996 CTACATTTGAGGCTTCCAAAGTGAAGGATAGGCTGAGAGATGCGACATTTGGGCTCTGAGGT 1055
Qy 822 GTCCCTGTCTGGGCACTGATCAAGAAACATCATCTCTCATTTGCTGCTGCTGCTGCTGCT 881
Db 1056 GACACCGGCTAGGAGTGGTGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
Qy 882 AAGGGTTCATGCGCAGCAGCTCGGGG---CCAGCTCAAGCGCGGAGCTCATAGGCTCCG 938
Db 1116 AAGGGAGCGCTGGAGCGCCACTCAGATTTGGCCACCCCTGAAACCCCACTCTGCTGGG 1175
Qy 939 AGATATGAGGCCCTGATTGCAAAACCGAGCTAGCTTCTCTCTCTCTCTCTCTCTCTCTCT 998
Db 1176 AGATCTGAGGGCTCTGGCAGAAATCTTAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 1235
Qy 999 TGTCTCTTTCGCAACCGGGGCGCTGGGCACTGATGATCCGCTCTCTCTCTCTCTCTCTCTCT 1058
Db 1236 TGTGGCTTTGTCAAGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1295

Db 614 GATCGTGTCTCCACCTGCTGGGTGATGCTGGGAGCCCTACCTCATTTGGCCGTGATCTC 555
Qy 1476 AGACCTGATCCGCGAGACACTAAGGACTCCCGCTCTGGGAGTTCTCTGAGCTGGGCTA 1535
Db 554 TGACCGCTGCGCGGAACCTGGCCCCCTCTCTTCTTGTCCGAGTTCCGGGCTCTGCAGTT 495
Qy 1536 CGCGCTCATGCTCTGCCCTTTCTGCTGTGCTCTGGGCGGATGTTCTCTGCGCACTGC 1595
Db 494 CTCGCTCATGCTCTGCGCGTTTGTGGGCACTGGGCGGCGACAGCTTCTCTGGGACGCG 435
Qy 1596 GCTCTTCTCTGATGAGCGGCGGCGGCTGAGCAGCACCTGGGG 1642
Db 434 CATCTTCAATGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 388

RESULT 13

US-09-809-391-245
; Sequence 245, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 245
; LENGTH: 2201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-809-391-245

Query Match 26.28; Score 542.8; DB 10; Length 2201;
Best Local Similarity 64.38; Pred. No. 1.2e-127;
Matches 877; Conservative 1; Mismatches 478; Indels 8; Gaps 4;

Qy 285 CCCCGCGCTCAGCAGCCCAACCGCGCAGCTTGGGCGCGGCGGCGGCGGCGGCGGCGGCGG 344
Db 416 CCAGGARGGCTGAGCGGATCAMCGGCTGTCTCCCGGCGGCTTGGCTCTCATAGTGGC 475
Qy 345 CATCTCAGCTTGGGCAACGTGTCTAACTACTGGACAGGTACACCGTGGCAGCGGCTCT 404
Db 476 GGTGCTGTGTACATCAATCTCTGAACTACATGACCGCTTACCGTGGCTGGCGTCT 535
Qy 405 TCTGGACATCCAGCAGCACTTGGGCTCAAGGACCGAGCGCGGCTGTGAGTCAGT 464
Db 536 TCCGACATCGAGGAGTCTTCAACATCGGGGACAGTGTCTGGGCTCATCCAGACCGT 595
Qy 465 GTTCATCTGTAGCTTTCATGTGGCTGCGCCCATCTTCCGCTACTCTGGGCGACCGCTTCAA 524
Db 596 GTTCATCTCAGTTACATGTGTGGACACTGTGTGGTGTACTCTGGGTGACAGGTACAA 655
Qy 525 CAGGAAGTGTCTCAGCTCGCGCAATTTCTTCTGTGTGGCGGCTCACTTCTCCAGCTC 584
Db 656 TCGGAAGTATCTCATGTGCGGGGCAATGCTTCTGTCTCTGTCTGTGTGACACTGGGGTCA 715
Qy 585 CTTCATTTCCCGCAGTACTTCTGGCTGTGCTGTCTCTGCTCTCCGGGGGCTGTGGGCTCGG 644
Db 716 CTTCATTTCCCGGAGAGCAATTTCTGGCTGTCTCTCTGCTCTGACCGGGGCTGTGGGGT 775
Qy 645 GGAGGCGAGTACTCTCAACCATCCGCCCACTATCTTGGGACCTCTTCAACCAAGAACAC 704
Db 776 GGAGGCGAGTATTTCAACCATCCGCCCACTCTCATTTGGGACCTCTTGTGGCGGACCA 835
Qy 705 GGTACGCTCAGCTGTGCTCTTCTTACTTTCGCACTCCCACTGGGAGTGGCTGGGCTA 764
Db 836 GCGGAC-CGAGTCTCAGCATCTTCTTACTTTTGCAATTCGGTGGGAGTGGTCTGGGCTA 894
Qy 765 CATTAATGGCTCAGCGGTGAGCAGGCGGCGGAGCTGGCACTGGGCACTGGGGGTGTC 824
Db 895 CATTCAGGCTCCAAAGTGAAGATATGGCTGGAGACTGGCACTGGGCTCTGAGGGGTGAC 954

Qy 825 CCCTGTCTCTGGGCAATGATCAGGAACACTCATCTCTATTCTGTGTCAGCCACTATAAAG 884
Db 955 ACCGGTCTAGGAGTGTGTGGCGTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1014
Qy 885 GGTGATGCGGACAGCTCGGGGA---CAGCTCAAGGCCCGGACCTCATGGCTCCGAGA 941
Db 1015 GGGAGCGGTGGAGCGCACTCAGATTTGGCACCCTGAACCCCACTCTGTGTGGGCGA 1074
Qy 942 TATGAAGCGCTGATTTCGAACCGCAGCTAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1001
Db 1075 TCTGAGGGCTCTGGCAAGAAATCCTAGTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1134
Qy 1002 CTCCTTCCGACCGGGGCTCTGGGCAATGTGATCCCGCTCTACCTGACCCGCGCCCAAGT 1061
Db 1135 GGCCTTTGTCAAGCGGCTCTCTGGCTCTGTGGGCTCCGGCAATCTCTGTCTGTCTCTGTCT 1194
Qy 1062 TGTGCA---GAAAGACAGAGAGAGTGAACAGAGCCCGCTCTGTGGGCGCAAGAGCAGCT 1118
Db 1195 GGTCTTTGGGAGAGACCCCACTGCTTCCCGGAGACTCTCTGTCTCTCTCTCTCTCTCTCT 1254
Qy 1119 CATCTTTGGGCGCATCACTGCTTTACGGGATTTCTGGGCGTGTGTACGGGGGCGAGGAGC 1178
Db 1255 CATCTTTGGGACTCATCACTGCTTCCCGGAGTCTGGGCTGTGGGCTGTGGGCTGTGGAGAT 1314
Qy 1179 CAGCGCTGTGTGCGGCTGAAAGACCCAGCGGCGGCGGCGGCTGTGTGTGTGTGTGTGTGT 1238
Db 1315 CAGCGCGGCTGCGGCACTCCAAACCCCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1374
Qy 1239 GCTGGGCTGTGCATCTTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
Db 1375 CTTGGGCTGTGCACCTTCT 1434
Qy 1299 AGCTATATCTGTATCTTCTCGGGGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1358
Db 1435 CACTTATATTTTATCTTCAATGGAGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1494
Qy 1359 AGACATCTCTCATGTAGCTGT 1418
Db 1495 CGACATCTCTGT 1554
Qy 1419 CTTTCACTCTCCACTGTGTGGGAGCGCGGAGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1478
Db 1555 CGTGTCTGTCCACTGTGTGGT 1614
Qy 1479 CTTGATCGGCGCAGCACTAAGGACTCTCCGCTCTGGGAGTCTCTGAGCTCTGGGCTAGCG 1538
Db 1615 CCGCTTGGCGCGAACTGGGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1674
Qy 1539 GCTCATGCTCTGCGCTTCTGT 1598
Db 1675 GCTCATGCTCTGCGCTTGTGTGGGCACTGGGCGGCGGCACT-TTCTTGGGCGGCGGCACT 1733
Qy 1599 CTTTCTTGTGAGCGACCGCGCCAGGGCTGAGCAGCACTGGGGG 1642
Db 1734 CTTTCAATGAGCGGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1777

RESULT 14

US-09-882-171-245
; Sequence 245, Application US/09882171
; Publication No. US20030175858A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/882,171
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/809,391
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493

| Result No. | Score | Query | | | DB | ID | Description |
|------------|-------|-------|--------|-----|-----------|------------|-------------|
| | | Match | Length | No. | | | |
| 1 | 666 | 32.2 | 668 | 4 | BM722582 | UI-R-E00- | |
| 2 | 645.8 | 31.2 | 958 | 4 | B41710266 | B22582 | |
| 3 | 630.8 | 30.5 | 924 | 4 | B1414356 | UI-B-EO0- | |
| 4 | 596.6 | 28.8 | 828 | 7 | CN528616 | B22582 | |
| 5 | 575.8 | 27.8 | 772 | 7 | CK479120 | UI-N-HQ0- | |
| 6 | 557.8 | 26.9 | 837 | 4 | B1159625 | AGENCOUR | |
| 7 | 552.8 | 26.7 | 814 | 3 | AF370423 | 602920091 | |
| 8 | 524.2 | 25.3 | 543 | 4 | BM147219 | TCARAP1010 | |
| 9 | 512.4 | 24.7 | 1075 | 2 | B2285645 | BM147219 | |
| 10 | 509.8 | 24.6 | 687 | 4 | BG243478 | B2285645 | |
| 11 | 506.6 | 24.5 | 524 | 7 | CR556576 | 601097128 | |
| 12 | 502.6 | 24.3 | 815 | 7 | CO401999 | BG243478 | |
| 13 | 498.8 | 24.1 | 724 | 4 | BG974364 | CR556576 | |
| 14 | 495.2 | 23.9 | 1587 | 9 | AY403401 | DKF2p459J | |
| 15 | 483 | 23.3 | 579 | 7 | CN666419 | AGENCOUR | |
| 16 | 474.6 | 22.9 | 1587 | 9 | AY403403 | 602844132 | |
| 17 | 471.2 | 22.8 | 549 | 2 | B233986 | AY403401 | |
| 18 | 470.8 | 22.7 | 1430 | 3 | CR591379 | AY403403 | |
| 19 | 470 | 22.7 | 1615 | 3 | CR596645 | Mus muscu | |
| 20 | 468.8 | 22.6 | 771 | 2 | BF099708 | B233986 | |
| 21 | 467.4 | 22.6 | 989 | 4 | BX176211 | CR591379 | |
| 22 | 453 | 21.9 | 518 | 5 | B2280500 | CR596645 | |
| 23 | 442.6 | 21.4 | 668 | 2 | B3377721 | full-len | |
| 24 | 442 | 21.3 | 840 | 4 | B1765222 | full-len | |

```

FEATURES
    source
        seq primer: M13 reverse.
        Location/Qualifiers
            1..668
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="UI-E-E00-ahy-f-02-0-UI"
                /tissue_type="fetal eye"
                /dev_stage="fetal"
                /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                /clone_lib="UI-E-E00"
                /note="Organ: eye; Vector: pTT3-Pac (Pharmacia) with a
                    modified polylinker; Site_1: EcoR I; Site_2: Not I;
                    UI-E-E00 is a cDNA library containing the following
                    tissue(s): fetal eye. The library was constructed

```

according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGGATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 32.2%; Score 666; DB 4; Length 668;
Best Local Similarity 99.7%; Pred. No. 5.7e-125;
Matches 666; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 475 AGCTTCATGCTGCTGCCCCCATCTTCGGCTACCTCGGCGACCGCTTCAACAGGAAGGTG 534
DB 1 AGCTTCATGCTGCTGCCCCCATCTTCGGCTACCTCGGCGACCGCTTCAACAGGAAGGTG 60

QY 535 ATTCTCAGTGGCGCATTTCTTCTGGTTCGGCGCTCACCTTCTCAGACTCCTTCATTCCTCC 594
DB 61 ATTCTCAGTGGCGCATTTCTTCTGGTTCGGCGCTCACCTTCTCAGACTCCTTCATTCCTCC 120

QY 595 GACGAGTACTTCTGGCTGCTGCTCTCTCCGGGGGCTGGTGGCATCGGGGAGGCCAGC 554
DB 121 GACGAGTACTTCTGGCTGCTGCTCTCTCCGGGGGCTGGTGGCATCGGGGAGGCCAGC 180

QY 655 TACTCCACCATCGCCCCCACTATCATTTGGCGACCTCTTCAACAGGAACACGGGTACGCTC 714
DB 181 TACTCCACCATCGCCCCCACTATCATTTGGCGACCTCTTCAACAGGAACACGGGTACGCTC 240

QY 715 ATGCTGTCCGCTTCTTCTTCTGCTTCCCATTCCTGCGGAGTGGCTTGGGCTTACATTAATGCG 774
DB 241 ATGCTGTCCGCTTCTTCTTCTGCTTCCCATTCCTGCGGAGTGGCTTGGGCTTACATTAATGCG 300

QY 775 TCCAGGCTGAACGACGACGCGGAGACTGGCACTGGGCACTGGGCTGTCCTGTCCTGTCCTG 834
DB 301 TCCAGGCTGAACGACGACGCGGAGACTGGCACTGGGCACTGGGCTGTCCTGTCCTGTCCTG 360

QY 835 GGATGATCACAGGAACACTCATCTCTCTGCTCCAGCCACTAAAGGGGTCTATGCC 894
DB 361 GGATGATCACAGGAACACTCATCTCTCTGCTCCAGCCACTAAAGGGGTCTATGCC 420

QY 895 GACGAGCTCGGAGCAGCTCAAGGCCGAGACTCATGCTTATGGCTCGAGATATGAAGGCCCTG 954
DB 421 GACGAGCTCGGAGCAGCTCAAGGCCGAGACTCATGCTTATGGCTCGAGATATGAAGGCCCTG 480

QY 955 ATTCGAAACCGGAGCTACGCTTCTCTCTGCTGGCCACGCTCGGCTGCTCTCTGCTGCTG 1014
DB 481 ATTCGAAACCGGAGCTACGCTTCTCTCTGCTGGCCACGCTCGGCTGCTCTCTGCTGCTG 540

QY 1015 GGGGCCCTCGGAGTGTGGATTCGGCTCTACCTGCACCGCGGCCCAAGTGTGCAAGAGACA 1074
DB 541 GGGGCCCTCGGAGTGTGGATTCGGCTCTACCTGCACCGCGGCCCAAGTGTGCAAGAGACA 600

QY 1075 GCAGAGACGTGCAACAGCCGCTCTGCGGCGCAAGGACAGCTCATCTTTGGGGCCATC 1134
DB 601 GCAGAGACGTGCAACAGCCGCTCTGCGGCGCAAGGACAGCTCATCTTTGGGGCCATC 660

QY 1135 ACCTGCTT 1142
DB 661 ACCTGCTT 668

RESULT 2
BI410266
LOCUS 602966259F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5121422 5',
DEFINITION mRNA sequence.
ACCESSION BI410266
VERSION BI410266.1 GI:15171189

KEYWORDS

SOURCE
ORGANISM

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 958)

REFERENCE

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM11296 row: f column: 15

High quality sequence start: 13

High quality sequence stop: 850.

Location/Qualifiers

1..958

/organism="Mus musculus"

/mol_type="mRNA"

/strain="CZECH II"

/db_xref="taxon:10090"

/clone="IMAGE:5121422"

/tissue_type="pooled lung tumors"

/lab_host="DH10B (phage-resistant)"

/clone_libs="NCI CGAP Lu33"

/note="Organ: lung; Vector: p7T73D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTT 3'].

Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 31.2%; Score 646.8; DB 4; Length 958;
Best Local Similarity 85.2%; Pred. No. 4.7e-121;
Matches 745; Conservative 0; Mismatches 121; Indels 2; Gaps 2;

QY 826 CCTGTCTCGGGCATGATCACAGGAACACTCATCTCTCTGCTCCAGCCACTAAAGG 885
DB 6 CCCCACTGGGCATGATCACAGGAACACTCATCTCTCTGCTCCAGCCACTAAAGAG 65

QY 886 GGTCTATGCGACCACTCGGGACACAGTCAAGGGCCGGACCTCATGGTCCGAGATATG 945
DB 66 GGCATGTGTCAACTTGGGGGCGAGTCAAGAGCAGGACCTCTCTGCTTCGAGACATG 125

QY 946 AAGGCCCTGATTCGAAACCGGAGTACGTTCTCTCTCTGCTCCAGCCACTGCTCTCC 1005
DB 126 AAGGCCCTGATCCGAAACCGGAGTACGTTCTCTCTCTGCTCCAGCCACTGCTCTCC 185

QY 1006 TTTCGCACGGGGCCCTCGGGCATGTGGATTCGGCTCTTACCTGCACCGCCCAAGTTGTG 1065
DB 186 TTTCGCACAGGGGCCCTCGGGCATGTGGATTCGGCTCTTACCTGCACCGCTCAAGTTGTA 245

QY 1066 CAGAGACAGCAGAGACGTGCAACAGCCGCCCTGTGGGGCCCAAGAGACCTCATCTTT 1125
DB 246 CAAAAGACAGCAGAGACTTGCAACAGCCGCCCTGTGGAGCCCAAGAGACCTCATCTTT 305

QY 1126 GGGGCCATCAGCTGCTTTTACGGATTCTTGGCGGTGTTCCACGGGGCAGGAGCCAGCGC 1185
DB 306 GGAGGCATTTACTGCTTTTACTGCTTCTTCTGCGGCGTAGTCACGGGGCCGGAGCCATCGT 365

QY 1186 TGGTCCGCTCGAAGACCCAGCGGGCCGACCCACTGCTGGTGTGTGCGGTGGGCATGCTGGGC 1245

```

|||||
366 TGGTGGCCCTCGGACCTCAACGCTGCTACCCATGGTGTGTGTGGGATGCTGGGA 425
|||||
1246 TCTGCCATCTTCATCTGCTGATCTTCGTGCTGCCAAGACAGCATCGTAGGACCTAT 1305
|||||
426 TCGCCCATCTTCATCTGCTCATCTTTGTGTGCTGCCAAGACAGCATGTGAGGCGCTAT 485
|||||
1306 ATCTGTATCTTCGTGCGGAGACGCTGCTGTTTTCTTAACCTGGGCATCATCTGCAGACATC 1365
|||||
486 ATCTGCATTTTTGTGGAGAGACATCTGCTGTTTTCTAACTGGGCATCATCTGCAGACATC 545
|||||
1366 CTATGTACGTGGTTCATCCCAACGCGCGCCACTGCGGCGCTTGCAGAGCTTCACC 1425
|||||
546 CTATGTATGGTTCATCCCACTCGCGAGCCACTGCTGCGCCCTGCAGAGCTTCACC 605
|||||
1426 TCCCACTGCTGGGAGCGCGGAGCCCTTACCTCATTTGCTTTATCTCAGACCTGATC 1485
|||||
606 TCCCATCTGCTGGGAGCGCTGGAAGCCCTTACCTCATTTGCTTTATCTCGACCTATC 665
|||||
1486 GCGCAGACCTAAGGACTCTCCCGCTCTGGAGTTCTTGAGCCTGGGCTAGCGCTCATG 1545
|||||
666 GCGCAGACACCAAGGACTCTCCCGCTCTGGAGTTCTTGAGCCTGGGCTATGCCCTCATG 725
|||||
1546 CTCTGCCCTTCGTGCTGGTCTCTGGCGGCA-TGTTCTTCTCGCACTGGGCTCTTCTT 1604
|||||
726 CTGTGCCCTTTTGTGCTGGTCTCTGGTGGCATTTGTTCTTCTTCCCACTGCTCTTCTT 785
|||||
1605 CGTCAGCGACCGCCAGGCTGAGCAGCACTCGGGGAGAGAGCGGGGGGTCAGGGT 1664
|||||
786 CCTCAGCGACCGTGCCA-GGCTGAGCAGCAGGTGACACGCTGTGTATGCTCTCGCATCC 844
|||||
1665 GGTGCATCAGCGGGCGCGGCCCGGCACCTGCT 1698
845 GGAAGTCTCAGGGGGTGGCGCTGGGACAGCACT 878

```

```

RESULT 3
BI414356 924 bp mRNA linear EST 14-AUG-2001
LOCUS 602990376F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5146410 5',
DEFINITION mRNA sequence.
ACCESSION BI414356
VERSION BI414356.1 GI:15175279
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 924)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11361 row: 9 column: 19
High quality sequence start: 10
High quality sequence stop: 824.
Location/Qualifiers
1. .924
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5146410"
/tissue_type="pooled lung tumors"

```

FEATURES

source

```

/lab_host="DH10B (phage-resistant)"
/clone_lib="NCI_CGAP_Lu33"
/note="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'.
TGTTACCAATCTGAAGTGAGGCGCGCTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified p773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

```

ORIGIN

```

Query Match 30.5%; Score 630.8; DB 4; Length 924;
Best Local Similarity 86.5%; Pred. No. 8.4e-118;
Matches 754; Conservative 0; Mismatches 112; Indels 6; Gaps 5;

Qy 755 GCCTGGGCTACATTACTTGGCT--CCAGCGTGAAGCAGCGAGCCGGAGACTGGCACTGGGC 812
Db 8 GCGTGGGCTATATCACAGGCTTCAGCGTTGAAGCAGCGAGCTGGAGACTGGCACTGGGC 67
Qy 813 ATTGCGGGTCTCCCTGCTCTGGGCATGATCAGAGGAACATCATCTCTCATCTTGGTCCC 872
Db 68 CCTGCGGGTTTCCCGCTCTGGGCATGATCAGAGGAACATCATCTCTCATCTTGGTCCC 127
Qy 873 AGCCACTAAAGGGGTCTATGCCAGCAGCTCGGGGACAGCTCAAGCCCGGACCTCATG 932
Db 128 AGCCACTAAGAGGGCCATGCTGATCACTTGGGGGAGCTCAAGACGAGACCTCTCTG 187
Qy 933 GCTCCGAGATATGAAGCCCTGATTTCGAAACCGCAGCTACGTTCTTCTCTCCCTGGCCAC 992
Db 188 GCTTCGAGACATGAAGGCCCTGATCCGAAACCGCAGTTACGTTCTTCTCTCCCTGGCCAC 247
Qy 993 GTGGGCTGCTCTCTTCCGACGGGGGCTTGGGCATGTGATCCGCTCTACCTGCACCG 1052
Db 248 ATCCGCTGTCTCTTCCGACAGGGGCTTGGGCATGTGATTTCCCTCTATCTTCCACG 307
Qy 1053 CGCCCAAGTTGTGCAGAAAGACAGCAGAGAGCTGCAACAGCCCGCTGTGGGGCCAGGA 1112
Db 308 TGTCTAAGTTGTACAAAGACAGCAGAGACTTGCACAGCCGCTGTGGAGCAAGA 367
Qy 1113 CAGCCTCATCTTTGGGCGCATCACTGCTTTAGCGGATTTCTGGGCGTGTGTCAGGGGCG 1172
Db 368 CAGCCTCATCTTTGGAGCCATTACCTGCTTTACTGGCTTCTGGGCGTAGTCACGGGCGC 427
Qy 1173 AGAGACCAACGCTGTGGCGCTGAAAGACCCAGCGGCGCGACCACTGTTGTGTGCGGT 1232
Db 428 GGGAGCCACTCGTTGGTGGCGCTTGGGACTCAACGCTGCTGAGCCATTGGTGTGTGCTGT 487
Qy 1233 GGGCATGCTGGGCTCTGCCATCTTCACTGCGCTGATCTTCTGCTGCTGCCATCTTGT 1292
Db 488 GGGCATGCTGGGATCGGCCATCTTCACTGCTGCCATCTTGTGTGCTGCCAAGACAGCAT 547
Qy 1293 CGTAGGAGCCATATCTGTATCTTCTGCGGGAGAGCGCTGCTGTTTTTCTAACTGGGCGCAT 1352
Db 548 TGTAGGCGCTTATCTGCAAT-TTTTGTGGAGAGACACTGCTGTTTTTCTAACTGGGCGCAT 606
Qy 1353 CACTGCAGACATCTCATGTATGTCATCCACGCGGCGCGCACTGCGGTGGGCGCTT 1412
Db 607 CACTGCAGACATCTCATGTATGTCATCCCACTCGCGGAGCCACTGCTGTGGGCGCTT 666
Qy 1413 GCAGAGCTTCACCTCCCACTGCTGGGAGCGCGGGAGCCCTTACCTCATCTGGCTTTAT 1472
Db 667 GCAGAGCTTCACCTCCCACTGCTGGGAGCGCTGGAAGCCCTTACCTCATTTGGC-TTAT 725
Qy 1473 CTGAGACCTGATCCGCGCAGACACTAAGGACTCCCGCTCTGGGAGTTCTTGAAGCCCTGGG 1532
Db 726 CTGAGACCTTAATCCG-CAGAGCAACCAAGNACTCCCGCTCTGGAGAGTCCCTGAGCGCTGG 784
Qy 1533 CTACGCGCTCATGCTCTGCCCTTTTCTGTGTGCTCTGGGCGGCAATG-TTCTTCTCGCCA 1591
Db 785 CTATGCCCTCATGCTGTGCCCTTTTGTGTGTGCTCTGGGAGGCAATGTTTCTTCTTGGCCA 844

```


DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM15012 row: e column: 20
High quality sequence stop: 692.
Location/Qualifiers
1. .772
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7123774"
/tissue_type="lung, pooled"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 232"
/note="Organ: lung; Vector: pExpress-1; Site 1: EcoRV;
Site 2: NotI; RNA obtained from pooled lung tissue from a
mix of male and female animals at 8 wk old. Tissues were a
snap-frozen and kept at -80C for two days before RNA
extraction and purification (TRI-reagent method). cDNA was
5'-primed using oligo-dT primer:
5'-pGACATGTTCTAGATCGAGCGCGGCCCT(T)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb
resulted in an average insert size of 2.3 kb. This primary
library is normalized (non-normalized primary library is
NIH MGC 231) and was constructed by Express Genomics
(Frederick, MD). Note: this is a NIH MGC library."

ORIGIN

| Query Match | 27.8%; Score 575.8; DB 7; Length 772; |
|---------------------------|--|
| Best Local Similarity | 87.8%; Pred. No. 1.3e-106; |
| Matches 628; Conservative | 0; Mismatches 87; Indels 0; Gaps 0; |
| QY | 802 TGGCACTGGGCATTGCGGGTGTCCTCGGCATGATCAGAGAACACTCATCTCTC 861 |
| DB | 1 TGGCAATGGGCCTTCGGGGTTTCCCTTGTCCTGGGCATGATCAGAGAACACTCATCTCTC 60 |
| QY | 862 ATTCTGGTCCAGCCACTAAAGGGGTCAATGCCACAGCTCGGGGACCAAGCTCAAGGCC 921 |
| DB | 61 ATCTTGTTCCAGCCACTAAGAGAGGGCATGCTGATCAACTGGGGGGCAGCTCAAAGCA 120 |
| QY | 922 CGGACCTCATGGCTCCGAGATATGAGGCCCTGATTCGAAACCGCAGCTACGTCCTCTCC 981 |
| DB | 121 CGGACCTCTTGGCTCCGAGACATGAAGGCCCTGATTCGGAACCGCAGTTCACGTCCTCTCC 180 |
| QY | 982 TCCTTGCCACAGCTCGGTGTCTCTTCCTCGCCACGGGGGCCCTGGGCGATGTGGATCCCGCTC 1041 |
| DB | 181 TCCTTGCCACATCCGCTGTGTCCTTCGCCACAGGAGCCCTGGGCATGTGGATTCCTCTC 240 |
| QY | 1042 TACCTGCACCGCGCCCAAGTTGTGCGAAGACAGCAGAGAGCTGCAACAGCCCGCCCTGT 1101 |
| DB | 241 TATCTGCACCGTCTCAAGTTGTACAAAGACAGCAGAGACTTTCGAACAGCCCGCCCTGT 300 |
| QY | 1102 GGGGCCAAGGACAGCCTCATCTTTTGGGGCCATCACTGCTTTACGGGATTTCTGGGCGCTG 1161 |
| DB | 301 GGAGCCAAAGACAGCCTCATCTTTGGGGCCATTAACCTGCTTTACTGCTTTCTGGGCGTA 360 |
| QY | 1162 GTCACGGGGCAGGAGCCACGCGCTGTGTCGCTGAAGACCCAGCGGGCCGACCACTG 1221 |
| DB | 361 GTCACAGGCGGGAGGCACTCGCTGGTGGCGCTTTCGCGACTCAGCGCGCTGACCCCTTG 420 |
| QY | 1222 GTGTGTGCGGTGGGCATGCTGGGCTTCGCATCTTTCATCTGCTGATCTTCGTGGGCTGCC 1281 |
| DB | 421 GTGTGGCCCGTGGGCATGCTGGGATTCGCCATCTTTCATCTGCTCATCTTTGTGGCTGCC 480 |
| QY | 1282 AAGACGAGCATCGTAGGAGCCTATATCTGTATCTTCGTGGGGAGACGCTGCTGTGTTTCT 1341 |
| DB | 481 AAGACGAGCATCGTGGGGCCATATCTGCATCTTTGTTGGCGAGACACTGCTGTGTTTCT 540 |
| QY | 1342 AACTGGGCCATCACTGACAGACATCTCTATGTAGCTGGGTATCCCAACGGCGCGCCACT 1401 |
| DB | 541 AACTGGGCCATCACTGACAGATCTCTCATGTATGTGGTCACTCCCACTCGACGAGCACT 600 |

| | | | |
|-----------------------|---|--|----------------------------|
| Qy | 1402 | GCCGTGCGCCCTTGAGAGCTTACCTCCCACTGCTGCTGGGGAGCCCGGAGCCCTTACCTC | 1461 |
| Db | 601 | GCTGTGCGCCCTTGAGAGCTTACCTCCCACTGCTGCTGGGGAGCGCTGGAAGCCCTTACCTC | 660 |
| Qy | 1462 | ATTGGCTTTATCTCAGACCTGATCCGCCAGAGACACTAAGGACTCCCGGCTCTGGG | 1516 |
| Db | 661 | ATTGGCTTTATCTCGGACCTTAATCCGCCAGAGACAATGACTCCCACTCTGGGAG | 715 |
| RESULT 6 | | | |
| BI159625 | | | |
| LOCUS | 6029200391F1 NCI_CGAP_Mam3 | Mus musculus cDNA clone IMAGE:5060351 | 5', linear EST 05-JUL-2001 |
| DEFINITION | mus musculus (house mouse) | | |
| ACCESSION | BI159625 | | |
| VERSION | BI159625.1 | GI:14619626 | |
| KEYWORDS | EST | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | 1 | (bases 1 to 837) | |
| AUTHORS | NIH-MGC | http://mgc.nci.nih.gov/ | |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) | | |
| JOURNAL | Unpublished (1999) | | |
| COMMENT | Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D. cDNA Library Preparation: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLM1163 row: e column: 24 High quality sequence start: 62 High quality sequence stop: 837. | | |
| FEATURES | | | |
| source | 1. .837 | | |
| | Location/Qualifiers | | |
| | /organism="Mus musculus" | | |
| | /mol_type="mRNA" | | |
| | /strain="129,C57BL/6J,FVB/N" | | |
| | /db_xref="taxon:10090" | | |
| | /clone="IMAGE:5060351" | | |
| | /issue_type="tumor, gross tissue" | | |
| | /dev_stage="10 months" | | |
| | /lab_host="DH10B" | | |
| | /clone_lib="NCI_CGAP_Mam3" | | |
| | /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)." | | |
| ORIGIN | | | |
| Query Match | 26.9%; | Score 557.8; | DB 4; Length 837; |
| Best Local Similarity | 87.3%; | Pred No. 5.9e-103; | |
| Matches 666; | Conservative 0; | Mismatches 92; | Indels 5; Gaps 5; |
| Qy | 745 | CTGGGCGAGTGGCTCGGGCTACA-TTACTGGCTCCAGCGGTGAAGCAGGACCGCGAGACTG | 803 |
| Db | 79 | CNCGGGGGTGGNCTGGGCTATAGTCACAGGTTCCAGCGGTGAAGCAGGCGAGCTGGAGACTG | 138 |
| Qy | 804 | GCACTGGGCATTTGCGGGTGTCCCCGTCTCTGGGCATGATCAGAGAACTCATCTCTCAT | 863 |
| Db | 139 | GCAATTGGGCGCCCTCGGGTTTCCCCGCTCTGGGCATGATCAGAGAACTCATCTCTCAT | 198 |
| Qy | 864 | TCCTGGTCCAGCGACCTAAAGGGGTGATGSCCGACCGAGCTCGGGGACCGAGCTCAAGGCCG | 923 |
| Db | 199 | CTTGTGTCACAGCACTAAGAGGGCCATGCGGATCAACTTGGGGGGCAGCTCAAGACAG | 258 |
| Qy | 924 | GACCTCATGGCTCCGAGATATGAAGGGCCCTGATTGGAACCCGACGCTACGTTCTCTCTC | 983 |

Db 259 GACCTCTGGCTTCGAGACATGAAGCCCTGTATCCGAAACCGAGTTAGCTTCTCTC 318
Qy 984 CTTGGCCAGCTGGCTGTCTCTCCGACAGGGGCGCTTGGGCATGTGGATCCGCTCTA 1043
Db 319 CTTGGCCACATCGCTGTCTCTTCCGACAGGGGCGCTTGGGCATGTGGATCCGCTCTA 378
Qy 1044 CTTGCAACCGCGCCCAAGTTGTGCGAGAGACAGAGAGAGCTGCAACAGCCCGCTGTGG 1103
Db 379 TCTTACCGCTGTCAAGTTGTACAAAGACAGCAGAGACTTGCACAGCCCGCTGTGG 438
Qy 1104 GCGCAAGACAGCTCATCTTTGGGCGCATCACTGCTTTACGGGATTTCTGGGGTGT 1163
Db 439 AGCCAAAGACAGCTCATCTTTGGAGCATTACCTGCTTTACTGGCTTCTGGGGTAGT 498
Qy 1164 CACGGGGGCGAGGACGACGCTGTGGTGGCGCTGGAAGACCCAGCGGGCGACCCACTGCT 1223
Db 499 CACGGGGGCGAGGACGACCTCGTTGGTGGCGCTGCGGACTCAACGCTGACCCATTGCT 558
Qy 1224 GTGTGCGTGGGATGCTGGGCTGTGGCATCTTTCATCTGCTGATCTTGGTGGCGCA 1283
Db 559 GTGTGCTGTGGGATGCTGGGATCGGCATCTTTCATCTGCTCATCTTTGTGGCTGCCA 618
Qy 1284 GAGCAGCATGTAAGAGCTATATCTGTATCTTCTGCGGGGAGACGCTGCTGTTTCTAA 1343
Db 619 GACCAAGCATGTAGGGCGCTTATCTGCAAT-TTGTGTGGAGAGACATGCTGTTTCTAA 677
Qy 1344 CTTGGGCGCATCTGCAGACATCTCATGTACGTGTGTCATCCCAACGCGGCGCCACTGC 1403
Db 678 CTTGGGCGCATCTGCAGACATCTCATGTATGTGTGTCATCCCACTCGCGGAGCCACTGC 737
Qy 1404 CGTGGCCTTGCAGAGCTTCACTTCCACCTGCTGGGGGAGCGCGGAGCCCTTACTCAT 1463
Db 738 TGTGGACCTGCAGAGCTTCA-CTCCCATCTGCTGGGGGAGCGTGGAGGCCCCCTAACTCAT 796
Qy 1464 TGGCTTTATCTCAGACTCATCGCCAGAGACTCAAGGACTCC 1506
Db 797 TGG-TTTATCTCGACCTTATCCG-CAGAGCCAAAGGACTCC 837

RESULT 7
AF370423 2144 bp mRNA linear HTC 31-DEC-2003
LOCUS Homo sapiens PP2030 mRNA, complete cds.
DEFINITION AF370423
ACCESSION AF370423.1 GI:33341775
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Huang, Y., Zhou, X.M., Zhang, P.P., Jiang, H.Q., Qin, W.X., Zhao, X.T.,
Wan, D.F. and Gu, J.R.
TITLE Novel human cDNA clones with function of inhibiting cancer cell
growth
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2144)
AUTHORS Zhou, X.M., Zhang, P.P., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T.,
Wan, D.F. and Gu, J.R.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2001) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
200032, P. R. China
FEATURES
source Location/Qualifiers
1..2144
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
597..1859
/codon_start=1
/product="PP2030"
/protein_id="AAQ15259.1"
CDS

/db_xref="GI:33341776"
/translation="MVLAPVRGYLDRNRYLMCGGIAFWSLVTLGSSFFPGHFHFL
LLLRVLGVGEASYSLIAPTLIADLVADORSMLSFYFAIPVGSGLGYTAGSKVK
DMAGDHWALKRVTPGLGVAVLLFLVVRPPRGAVRHSDLPLNPPSWADLRALA
RNPFSVLSSGFTAVFVTSRLASWAFLLRSRVLTGTPFPFPGDSCCSSDSLIAG
LITCLTGLVGLGVEISRLRHSNPRADPLVATGLGSAPFLLSLUACARGSVIF
YLFITGTLTMMNNAIVADILLYVVIPTRRSTAEAFQIVLSHLGLDAGSPYLGLIS
DLRNNPPSPFLSEFRALQFLMLCAFVGLGAAFLGTAIFIBADRRRAQLHVQGLL
HEAGTDDRIVVPPQRGRSTRVPASVLI"
ORIGIN
Query Match 26.7%; Score 552.8; DB 3; Length 2144;
Best Local Similarity 64.3%; Pred. No. 6.5e-102;
Matches 879; Conservative 0; Mismatches 482; Indels 7; Gaps 3;
Qy 282 GGGGGCGGCGCTAGACGCCAAACCGGCAGCTTGGGCGCGGGGGGCGAGCGC 341
Db 397 GGAACAGAGGGGCTGCAGCGCATCACCGGCTGTCTCCCGGCCCTTCATAGT 456
Qy 342 CGCCATCTCAGCTTGGGCAACGCTCAACTTACCTGACAGGTACACCGTGGCAGGCT 401
Db 457 GCGGTGCTGTGTACATCATCTCTGAACTATACAGACCGCTTACCGTGGCTGGCT 516
Qy 402 CTTCTGGACATCCAGC-AGCACTTTGGGGTCAAGACCGAGCGCGGCGCTGTGCTAGT 460
Db 517 CTTTCCGACATCGAGCAAGTCTTCAACATCGGGGACAGTAGCTCTGGGCTCATCAGA 576
Qy 461 CAGTGTTCATGTAGTCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 520
Db 577 CCGTGTTCATCTCCAGTTACATGTTGGGACCTGTGTTGGCTACCTGGGTGACAGT 636
Qy 521 TCAACAGAGAGGTGATCTCAGCTCGGCAATTTCTTCTGTGTGGCGCGTCACTTCTCCA 580
Db 637 ACATCGAGATATCTATGTGCGGGGCAATGCTTCTGTGCTCCCTGTGTGACACTGGGT 696
Qy 581 GCTCTTTCATTTCCAGCAGTACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
Db 697 CATCTTTCATCCCGAGAGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 756
Qy 641 TCGGGAGCGGAGCTTCCACCATCGCCCGCCACATCATTTGGCGACCTTTCACCAAGA 700
Db 757 TCGGGAGGCGGAGTTATTTCCACCATCGCGCCACATCTCATTTGCCAGCTTCTTTGGGCG 816
Qy 701 ACACGCTGACGCTCATGTGTCCTGCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG 760
Db 817 ACCAGCGGAGCGGATGCTCAGCATCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTG 876
Qy 761 GCTACATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 820
Db 877 GCTACATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
Qy 821 TGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 880
Db 937 TGACACCGGCTTAGAGTGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996
Qy 881 AAAGGGGTATGCCGACCACTCGGGG-CCAGCTCAAGGCCCGGACCTCATGAGCTCC 937
Db 997 CAAAGGGGAGCGGTGGAGCGCCACTCAGATTTGCGCAACCCCTGAACCCCTGCTGGTGGG 1056
Qy 938 GAGATATGAAGCGCTGATTCGAAACCGCAGCTAGCTCTCTCTCTCTCTCTCTCTCTCTCT 997
Db 1057 CAGATCTGAGGGCTCTGGCAAGAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1116
Qy 998 CTGTCT 1057
Db 1117 CTGTGCTCTTGTACAGGGCTCCCTGCTCTGTGGGCTCGGCAATTCCTGCTGCTGCTGCT 1176
Qy 1058 AAGTTGTGCA---GAAGACAGCAGAGACGTGCAACAGCCCGGCTGCTGAGCCCGGCC 1114
Db 1177 GGTGTGCTCTGGGGAGAGCCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1236
Qy 1115 GCTCATCTTTGGGGGCGCATCAGCTGCTTTACGGGATTTCTGGGCGTGGTCAACGGGGGCG 1174


```
Db 1237 GTCTCATCTTTGGACTCATACCTGCTCCCTGACCGGAGTCTCGGTGTGGGCTTGGGTGG 1296
Qy 1175 GAGCACCGCTGGTGGCGCTTGAAGACCCAGCGGCGAGCCCACTGCTGTGTGCGGTGG 1234
Db 1297 AGATCAGCGCGGCTCGGCCACTCCAAACCCCGGGCTGATCCCTGGTCTGTGCCACTG 1356
Qy 1235 GCATGCTGGGCTCTGCCATCTTCACTGCTGATCTTGTGGTGGTGGTGGTGGTGGTGGTGG 1294
Db 1357 GCCTCCTGGGCTCTGCCATCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1416
Qy 1295 TAGGAGCCTATATCTGTATCTTCTGCTGGGAGAGCTGCTGTTCTTAACCTGGGCGCATCA 1354
Db 1417 TGGCCACTATATTTCTATCTTCACTTGAGAGAGACCTCTCTGTCATGAACTGGGCGATCG 1476
Qy 1355 CTGCAGACATCTCTCATGTACGTGTGTCATCCACGCGGCGGCGCACTGCCCTGGCCCTTGC 1414
Db 1477 TGGCGGACATCTGCTGTACGTGTGATCCCTACCGAGGCTCCACCGCGGAGGCTTCC 1536
Qy 1415 AGAGCTTCACTCCACCTGCTGGGGAGCGCGGAGCCCTCACTCATTTGGCTTTATCT 1474
Db 1537 AGATCGTGTCTGCTCCACCTGCTGGGTGATGCTGGGAGCCCTACCTCATTTGGCTGTATCT 1596
Qy 1475 CAGACCTGATCCGCGAGACACTAAGGACTCCCGCTCTGGAGTCTCTGAGCTGGGCT 1534
Db 1597 CTGACCGCTGCGCGGAACCTGGCCCCCTCTTCTTGTGCGAGTCCGGGCTCTGCAGT 1656
Qy 1535 ACGGCTCATCTCTGCGCTTCTGCTGGGTGCTGGGCGGATGTTCTTCTCGCCACTG 1594
Db 1657 TCTCGCTCATGCTCTGCGGCTTGTGGGCACTTGGGCGGCGAGCTTCTTCTGGGCACTG 1716
Qy 1595 CGCTCTTCTTCTGCTCAGGACCGCGCGGCTGAGAGACACTCTGGGG 1642
Db 1717 CCATCTTCTTATGAGCGGACCGCGCGGCGGACAGCTGCAGTGCAGG 1764

RESULT 8
BM147219 543 bp mRNA linear EST 30-NOV-2001
LOCUS TCAAP1Q10621 Pediatric acute myelogenous leukemia cell (FAB M1)
DEFINITION Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1062, mRNA
sequence.
ACCESSION BM147219
VERSION BM147219.1 GI:17166851
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R., Jr.,
Gunnaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.
Pediatric Leukemia cDNA Sequencing Project (2001)
Unpublished (2001)
Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@ccc.org
Seq primer: M13 primer.
FEATURES
source
1..543
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCAAP1062"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH10B"
/clone_lib="Pediatric acute myelogenous leukemia cell (FAB
```

```
M1) Baylor-HGSC project-TCAA"
/note=Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'-GGAGGACTCGAGCGCGGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
was primed with a BamHI-dC primer
[5'-AGAGAGCTCGATCCGCGCGCAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
```

ORIGIN

```
Query Match 25.3%; Score 524.2; DB 4; Length 543;
Best Local Similarity 98.9%; Pred. No. 4e-96;
Matches 537; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1037 CGCTCTACCTGACCGCGCCCAAGTTGTGAGAAGACAGACAGCTGCAACGCCGC 1096
Db 1 CGCTCTACCTGCAACCGCGCCCAAGTTGTGAGAAGACAGACAGCTGCAACGCCGC 60
Qy 1097 CCTGTGGGGCAAGACAGAGCTCATCTTTGGGGCCATCAGCTGTTACGGGATTTCTGG 1156
Db 61 CCTGTGGGGCAAGACAGAGCTCATCTTTGGGGCCATCAGCTGTTACGGGATTTCTGG 120
Qy 1157 GCGTGTGTACGGGGGCGAGGA-GCCACGCGCTGTGTGCGCCCTGAAAGACCCAGCGGCGCAC 1215
Db 121 GCGTGTGTACGGGGGCGAGGATTTTCACGCGCTGTGTGCGCCCTGAAAGACCCAGCGGCGCAC 180
Qy 1216 CCAGTGTGTGTGCGGTGGGATGCTGGGCTCTGCCATCTTTCATCTGCTGATCTTCTGTG 1275
Db 181 CCAGTGTGTGTGCGGTGGGATGCTGGGCTCTGCCATCTTTCATCTGCTGATCTTCTGTG 240
Qy 1276 GCTGCCAAGACAGACATCGTAGGAGCTATATCTGTATCTTCTGTCGGGAGACGCTGCTG 1335
Db 241 GCTGCCAAGACAGACATCGTAGGAGCTATATCTGTATCTTCTGTCGGGAGACGCTGCTG 300
Qy 1336 TTTTCTAACTGGGCGCATCACTGCAGACATCTCATGTACGTGTGTATCTCCACCGCGGCGC 1395
Db 301 TTTTCTAACTGGGCGCATCACTGCAGACATCTCATGTACGTGTGTATCTCCACCGCGGCGC 360
Qy 1396 GCACATGCGCTGGCTTGCAGAGCTTCACTCCACCTGCTGGGGAGCGCGGAGCGCC 1455
Db 361 GCACATGCGCTGGCTTGCAGAGCTTCACTCCACCTGCTGGGGAGCGCGGAGCGCC 420
Qy 1456 TACCTCATTTGGCTTTATCTCAGACCTGATCCGCGAGAGCACTAAGGACTCCCGCTCTGG 1515
Db 421 TACCTCATTTGGCTTTATCTCAGACCTGATCCGCGAGAGCACTAAGGACTCCCGCTCTGG 480
Qy 1516 GAGTTCCTGAGCTGGGCTACGGGCTCATGCTCTGCGCTTTTCTGCTGTGTGCTCTGGGCGGC 1575
Db 481 GAGTTCCTGAGCTGGGCTACGGGCTCATGCTCTGCGCTTTTCTGCTGTGTGCTCTGGGCGGC 540
Qy 1576 ATG 1578
Db 541 ATG 543

RESULT 9
BE285645 1075 bp mRNA linear EST 26-OCT-2000
LOCUS 601097128F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:3495966 5',
DEFINITION mRNA sequence.
ACCESSION BE285645
VERSION BE285645.1 GI:9163408
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
```

| | | | |
|---|------|--|------|
| Db | 481 | TTTGTGGCTGCCAAGACAGCAATGTAGCGGCTTATATCTGCATTTTGTTCGAGAGACA | 540 |
| Qy | 1330 | CTGCTCTTTTCTAACTGGGCCATCACTGCA-GACATCCCTCATGTACGTGGTCAATCCCCAC | 1388 |
| Db | 541 | CTGCTGTTTTCTAACTGGGCCATCACTGCAGGACATTTCTCATGTCTGTGGTCAATCCCCAC | 600 |
| Qy | 1389 | GGGGCGCGCCACTGCCGTGGCTTCGAGAGCTTACCTCCACCTGCTGGGGGAGCGCG | 1447 |
| Db | 601 | TCGGCGAGCGCACTGCTGTGGCCCTCAGAGGTT-AACTTCATCTGCTGGGGGA--CCT | 657 |
| Qy | 1448 | GGAGCCCTACCTCATTTGCTTTATCT-CAGACCTGATCCGCCAGAGCACTAAGGACTCC | 1506 |
| Db | 658 | GGAGCCCTAACTCATTTGGGTTATCTCGGACCTAAATCCGAGAGAGCCAGGATCCC | 717 |
| Qy | 1507 | CCGCTCTGG-GAGTTCTGAGCCTGGGCTACGGGCTCATGCTTCGCCCTTTGCTCGTGGT | 1565 |
| Db | 718 | CGTCTGGGAGCTCCCTCAGCCTGGGGTATGCCCTTGTGTGTGCCCTTTGTGCCGGGA | 777 |
| Qy | 1566 | CCTGGCGGCATGTTCTTCTCTCGC | 1589 |
| Db | 778 | CACGCGGGGAGCATTTGTTTCCC | 801 |
| RESULT 10 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| ACCESSION | | | |
| VERSION | | | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| FEATURES | | | |
| source | | | |
| ORIGIN | | | |
| Query Match | | | |
| Best Local Similarity | | | |
| Matches 594; Conservative | | | |
| 24.6%; Score 509.8; DB 4; Length 687; | | | |
| 89.3%; Pred. No. 3.4e-93; | | | |
| 0; Mismatches 67; Indels 4; Gaps 4 | | | |
| 960 AAACCGCAGTACGTTCTTCTCCTCGGCCACGTCGGCTGTCTCTTCGCCACGGGGC | | | |
| 101 | | | |

AUTHORS
TITLE NIH-MGC <http://mgs.nci.nih.gov/>
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical College of Wisconsin
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM15381 row: o column: 16
 High quality sequence stop: 672.

FEATURES
 source
 Location/Qualifiers
 1..815
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7320138"
 /sex="both"
 /tissue_type="Brain - Pooled from several tissues from one or more individuals"
 /lab_host="DH10B TONa"
 /clone_lib="NIH MGC 255"
 /notes="Organ: brain/CNS; Vector: pExpress-1; Site 1: EORV; Site 2: NotI; RNA obtained from brain tissue of 8 wk old animal. Tissues were snap-frozen and kept at -80C before RNA extraction and purification (Tri-reagent method). cDNA was primed using oligo-dT primer: 5'-pGACTAGTCTACATCGCGAGCGCGCC(T)25-3' and cloned into the EORV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 1.7 kb. This primary library is a normalized (primary library is NIH MGC 254) and was constructed by Express Genomics (Frederick, MD).
 Note: this is a NIH_MGC library"

ORIGIN
 Query Match 24.3%; Score 502.6; DB 7; Length 815;
 Best Local Similarity 89.4%; Pred. No. 1e-91;
 Matches 541; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

1030 TGGATCCCGCTTACTGCAACCGCGCCCAAGTTGTGCGAGAGACGAGAGCTGCAAC 1089
 1 TGGATCCCGCTTACTGCAACCGCTGCTCAAGTTGTACAAAGACAGCAGAGACTTGCAAC 60
 1090 AGCCCGCCCTGTGGGCGCAGGACAGCTCATCTTTGGGGCCATCAGCTGTTTACGGGA 1149
 61 AGCCCGCCCTGTGGGCGCAGGACAGCTCATCTTTGGGGCCATCAGCTGTTTACGGG 120
 1150 TTTCTGGGCGTGTGTCAGGGGCGGAGGACCGCGCTGTGCGCGCTGGAAGACCCAGCGG 1209
 121 TTTCTGGGCGTGTGTCAGGGGCGGAGGACCGCGCTGTGCGCGCTGTCGCACTCAGGCG 180
 1210 GCGGACCCACTGCTGTGTGCGGCGGATGCTGGGCTCTGCCATCTTCATCTGCTGATC 1269
 181 GCTGACCCCTTGTGTGCGGCGTGGGATGCTGGGATCTGCCATCTTCATCTGCTCATC 240
 1270 TTGCTGCTGCCAGACAGCATCGTAGGAGCCTATATCTGTATCTTCGTCGGGGAGAGC 1329
 241 TTTGTGGCTGCCAAGACAGCATCGTGGGCGCCTATATCTGCACTCTTTGCGGAGACA 300
 1330 CTGCTGTTTCTTAAGTGGGCGCATCATCTGCAGACATCTCTATGACGTGGTGCATCCCCACG 1389
 301 CTGCTGTTTCTTAAGTGGGCGCATCATCTGCAGACATCTCTATGATGTGTTGTCATCCCCACT 360
 1390 CGGCGCGCCACTGCGGTGGCTTGCAGAGCTTCACCTCCACCTGCTGGGGGAGCGCGG 1449
 361 CGACGAGCCACTGCTGTGGCTTGCAGAGCTTCACTCCATCTGCTGGGGGAGCGCTGGA 420

1450 AGCCCTACCTCATTTGGCTTTTATCTCAGACCTGATCCCGCCAGAGCACTAAGGACTCCCCG 1509
 421 AGCCCTACCTCATTTGGCTTTTATCTCAGACCTAATCCGCGAGAGCACCAGGACTCCCCA 480
 1510 CTCTGGGAGTTCTCGAGCCTCGGCTACCGGCTCATGCTCTGCCCTTTGTCGTGTCCTG 1569
 481 CTCTGGGAGTTCTCGAGCCTCGGCTATGCCCTCATGCTGTGCCCTTTTGTGTGTCCTG 540
 1570 GCGGCGATGTTCTTCCTCGCAGCTCGGCTCTTCTGTCAGGACCGCGCCAGGCTGAG 1629
 541 GGTGCGATGTTCTTCCTTGGCCACTCTCTCTTCTTCAGGATCGTGGCCAGGCTGAG 600
 1630 CAGCA 1634
 601 CAGCA 605

RESULT 13
 BG974364
 LOCUS 602844132F1 NCI_CGAP_Mam4 Mus musculus cDNA clone linear EST 12-JUN-2001
 DEFINITION mRNA sequence.
 ACCESSION BG974364
 VERSION BG974364.1 GI:14362001
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 724)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10978 row: p column: 08
 High quality sequence stop: 719.

FEATURES
 source
 Location/Qualifiers
 1..724
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NMRI"
 /db_xref="taxon:10090"
 /clone="IMAGE:4979959"
 /tissue_type="tumor, gross tissue"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam4"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

ORIGIN
 Query Match 24.1%; Score 498.8; DB 4; Length 724;
 Best Local Similarity 85.5%; Pred. No. 5.9e-91;
 Matches 590; Conservative 0; Mismatches 97; Indels 3; Gaps 3;

691 TTCACCAAGAACACGCGTACGCTCATGCTGTC-CGTCTTCTACTTCGCCATCCACTGGG 749
 34 TTACCAAGNACACGCGACATGAGCTATCTAGTCTTCTATTGTCATCCCTGGG 93

QY 750 CAGTGGCTGGCTACATTACTGCTCCAGCTGAGCAGCAGCGGAGACTGGACTG 809
Db 94 CAGTGGCTGGCTATATACAGGTTCCAGCTGAAGCAGCAGCTGGAGACTGGCTTTG 153
QY 810 GGCATTGGGGTGTCCCTCG-TCTGGGCATGATCACAGGAACACTCATCTCATCTCG 868
Db 154 GGCCTTGGGGTGTCCCTCGATCTGGCATGATCACAGGACACTCATCTCATCTTTGG 213
QY 869 TCCAGGCACTAAAGGGGTGATCCGACAGCTCGGGACAGCTCAAGGCCCGGACCT 928
Db 214 TCCAGGCACTAAGAGGGCCATGCTGATCAACTTGGGGGCGAGCTCAAGGACGGACCT 273
QY 929 CATGGCTCCGAGATGATGAAGCCCTGATTCGAACCGCAGTACTGCTTCTTCCCTCG 988
Db 274 CTTGGCTTCGAGACATGAAGCCCTGATTCGAAACCGCAGTTACGTCTTCTTCCCTCG 333
QY 989 CCAGTGGCTGTCTCTTCCGACAGGGGCTTGGGCATGTGATCCCGCTTACTCTGC 1048
Db 334 CCATATCGCTGTCTCTTCCGACAGGGGCTTGGGCATGTGATTCCTCTTATCTTTC 393
QY 1049 ACCGCGCCCAAGTTGTGCAAGAGACAGCAGAGAGCTGCAACAGCCCGCTTGTGGGCCA 1108
Db 394 ACCGTGCTCAAGTTGTACAAAGACAGCAGAGACTTGCACAGCCCGCTTGTGGGCCA 453
QY 1109 AGGACAGCTCATCTTTGGGCCATACCTGCTTTTACGGGATTTCTGGGGTGTCTACGG 1168
Db 454 AAGACAGCTCATCTTTGGAGCCATTACCTGCTTTTACTGCTTCTGGGGTGTCTACGG 513
QY 1169 GGGCAGAGCCACGG-CTGGTGGCTGAGAGCCAGCGGGCGGACCCACTGGTGTGT 1227
Db 514 GCGGGGAGGACACTGATGGTGGCTGCGGCTGCGGACTCAAGCTGACCCACTAGGTGTGT 573
QY 1228 GCCGTGGGCATGCTGGGCTGTGCCATCTTCACTGCTGATCTTCTGGGCTGCCAAGAGC 1287
Db 574 GCTGTGGGCATGCTGGGATCCGCCATCTTCACTGCTCTCTTGTGGGCTGCCAAGACC 633
QY 1288 AGCATCTAGAGCTTATCTGATCTTCTGCTGGGAGAGCTGCTGTTTTCTAACTGG 1347
Db 634 AGCATAGTGGCGCTTATCTGATCTGATGAAGTGAAGGAGAGACACTGCTGATTACTAG 693
QY 1348 GCCATCACTGCAGACATCTCTCATGTAGTG 1377
Db 694 GCCATCACTGCAGACATCTCTCATGTATGTG 723

RESULT 14
AY403401
LOCUS
DEFINITION
Homo sapiens HCM1545 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY403401
VERSION
AY403401.1 GI:39759384
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1587)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL
Science 302 (5652), 1960-1963 (2003)
PUBMED
14671302
REFERENCE
2 (bases 1 to 1587)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
JOURNAL

Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1587
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1587
/locus_tag="HCM1545"
ORIGIN
Query Match 23.9%; Score 495.2; DB 9; Length 1587;
Best Local Similarity 58.4%; Pred. No. 3.3e-90;
Matches 799; Conservative 0; Mismatches 562; Indels 6; Gaps 2;
QY 282 GGGCCCGGGCTCAGCAGAGCCCAACCGGCAGCTTGGGCGGGGGGGGAGCGCGC 341
Db 126 GGAACAGAGAGGGGTGAGCGGCATCACCGGCTGTCTCCGCGCTCTCATAGT 185
QY 342 CGCATCTCTCAGCTTGGGCAACGCTGCTCAACTACCTGGAAGGTACACCGTGGCAGGCGT 401
Db 186 GCGGTGCTGTCTACATCAATCTCTGAATACATGACCGCTTACCGTGGTGGCGT 245
QY 402 CTTCTGGACATCAGCAGACACTTTGGGGTCAAGGACCGAGGGCCCGCTGTGCAATC 461
Db 246 CTTCCCGACATCGAGCAGTCTTCAACATCGGGGACAGTAGTCTTGGGCTCATCCAGAC 305
QY 462 AGTTTCATCTGATGCTTCAATGCTGCTGCCCCCATCTTGGCTACCTGGGCGACCGCTT 521
Db 306 CGTGTTCATCTCCAGTTACATGTTGGCACCTGTGTTGGCTACCTGGTGGACAGTA 365
QY 522 CAACAGAAAGTGATCTCAGCTGCGGCATTTTCTTCTGTGGCGCGTCACTTCTCCAG 581
Db 366 CAATCGAAGTACTCATGTGCGGGGCAATGCTTCTGTCTCTGTGACACTGGGTC 425
QY 582 CTCTTCATTTCCAGCAGTACTTCTGGCTGTGGTCTCTGTCCGGGGGTGTGGGAT 641
Db 426 ATCTTCATCCCGGAGAGCATTTCTGGCTGCTCTCTGACCCGGGCGCTGTGGGGT 485
QY 642 CGGGAGGCGAGCTACTCCACCATCGCCCCCACTATCTTGGGAGACCTCTTCCACAGAA 701
Db 486 CGGGAGGCGAGTATTTCACCATCGCGCCCACTCTCATTTGCCGACCTCTTTGTGGCGGA 545
QY 702 CACGCTACGCTCATGCTGCTTCTACTTCTGCTTCCCATCCACTGGCGAGTGGCTGGG 761
Db 546 CAGCGAGCGGAGTCTAGCATCTTCTACTTTGCCATTCGGTGGGAGTGTCTGGG 605
QY 762 CTACATTTACTGGCTCCAGCGTGAAGCAGGAGCGCGGAGACTGGCAGTGGGCAATGGCGGT 821
Db 606 CTACATTTGAGGCTCCAAAGTGAAGGATATGGCTGGAGACTGGCAGTGGGCTCTGAGGGT 665
QY 822 GTCCCTGTCTGGGCATGATCAGGAACACTCATCTCTATTTCTGTGTCTTCTGTGGGAGCGGCC 881
Db 666 GACACCGGGTCTAGGAGTGGTGGCGCTTCTGTGCTGTCTTCTGTGTGGGAGCGGCC 725
QY 882 AAGGGGTGATGCGGACGCTCGGGGA---CCAGTCAAGCGCGCGGACCTCATCGGCTCG 938
Db 726 AAGGGAGCGCTGGAGCGGCCACTCAGATTTGGCACCCCTGAACCCACCTCTGTGGTGGG 785
QY 939 AGATATGAAGGCCCTGATTCGAAACCGCAGTACTGTCTTCTCTCTCCCTGGCCAGCTGGC 998
Db 786 AGATCTGAGGGCTCTGCAAGAAATCCTAGTTTCTGTCTGTCTTCTCTCTGGGCTTCACTGC 845
QY 999 TGTCTCTTTCGACAGCGGGGCGCTTGGGCATGTGGATCCCGCTTACTCTGACCGCGGCCA 1058
Db 846 TGTGGCTTTGTCAACGGGCTCCCTGGCTCTGTGGGCTCCGCGCATTCCTGTGGGTTCCCG 905
QY 1059 AGTTGTGCA---GAAGACAGCAGAGCTGCAACAGCCCGCTGTGGGCGCCAGGACAG 1115
Db 906 CGTGGTCTTGGGGAGACCCACCTGCTTCCCGGAGACTCTCTGTCTTCTTCTCTGACAG 965
QY 1116 CTTCTATTTTGGGGCCATCACCTGTCTTTACGGGATTTCTGGGCGTGTGTACGGGGGCGAG 1175

Qy 1589 CCACTGGCGCTCTTCTTCGTCAAGCGACCGCGCCAGGGCTG 1627
Db 541 CCACTGGCTCTCTTCTTCGTCAAGCGACCGCGCCAGGGCTG 579

Search completed: October 23, 2004, 12:50:41
Job time : 6572 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 08:34:23 ; Search time 89 Seconds
(without alignments)
2281.357 Million cell updates/sec

Title: US-10-085-198-48

Perfect score: 2920

Sequence: 1 MMLECSAAGGABEEAD.....VHQGPQPGTALAHRVGAS 566

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-------------|
| 1 | 2920 | 100.0 | 566 | 5 | ADH48764 |
| 2 | 1978 | 67.7 | 398 | 4 | AAE04896 |
| 3 | 1604 | 54.9 | 312 | 7 | ADAE08326 |
| 4 | 1298 | 44.5 | 528 | 4 | AAE08326 |
| 5 | 1298 | 44.5 | 528 | 4 | AAE08326 |
| 6 | 1298 | 44.5 | 528 | 4 | AAE08326 |
| 7 | 1298 | 44.5 | 528 | 4 | AAE08326 |
| 8 | 1298 | 44.5 | 528 | 4 | AAE08326 |
| 9 | 1256.5 | 43.0 | 507 | 8 | ADAE08326 |
| 10 | 1250 | 42.8 | 581 | 4 | AAE08326 |
| 11 | 1222 | 41.8 | 455 | 5 | AAE08326 |
| 12 | 1176 | 40.3 | 244 | 3 | AAE08326 |
| 13 | 1167 | 40.0 | 476 | 4 | AAE08326 |
| 14 | 1167 | 40.0 | 476 | 4 | AAE08326 |
| 15 | 1115 | 38.2 | 420 | 7 | ADAE08326 |
| 16 | 1088 | 37.3 | 454 | 4 | AAE08326 |
| 17 | 1084 | 37.1 | 512 | 5 | AAE08326 |
| 18 | 1083 | 37.1 | 512 | 5 | AAE08326 |
| 19 | 1054 | 36.1 | 572 | 5 | AAE08326 |
| 20 | 1042 | 35.7 | 221 | 4 | AAE08326 |
| 21 | 1030 | 35.3 | 216 | 4 | AAE08326 |
| 22 | 964.5 | 33.0 | 605 | 4 | AAE08326 |
| 23 | 930.5 | 31.9 | 605 | 4 | AAE08326 |
| 24 | 855 | 29.3 | 531 | 7 | ADAE08326 |
| 25 | 855 | 29.3 | 531 | 8 | ADAE08326 |

| | | | | | | |
|----|-------|------|-----|---|----------|------------|
| 26 | 779 | 26.7 | 299 | 2 | AAW74805 | Human sec |
| 27 | 779 | 26.7 | 299 | 5 | ABG95256 | Human nov |
| 28 | 779 | 26.7 | 299 | 6 | ABO34450 | Region of |
| 29 | 779 | 26.7 | 299 | 7 | AD123111 | Novel hum |
| 30 | 779 | 26.7 | 299 | 8 | ADH74113 | Human sec |
| 31 | 744.5 | 25.5 | 552 | 4 | ABE62015 | Drosophila |
| 32 | 732 | 25.1 | 144 | 5 | ABE53205 | Human ORF |
| 33 | 672 | 23.0 | 424 | 2 | AAW74956 | Human sec |
| 34 | 672 | 23.0 | 424 | 5 | ABG95415 | Human nov |
| 35 | 672 | 23.0 | 424 | 6 | ABO34609 | Region of |
| 36 | 672 | 23.0 | 424 | 7 | AD123270 | Novel hum |
| 37 | 672 | 23.0 | 424 | 8 | ADH74272 | Human sec |
| 38 | 644 | 22.1 | 229 | 7 | ADM04605 | Human pro |
| 39 | 559 | 19.1 | 343 | 6 | ADAL1867 | Human nov |
| 40 | 553 | 18.9 | 304 | 6 | ADAL1625 | Human nov |
| 41 | 534 | 18.3 | 106 | 5 | ABE53122 | Human ORF |
| 42 | 533 | 18.3 | 248 | 7 | ADD78251 | Human CGD |
| 43 | 531 | 18.2 | 212 | 3 | AAE41388 | Human ORF |
| 44 | 465 | 15.9 | 256 | 6 | ABP75842 | Human sec |
| 45 | 459.5 | 15.7 | 204 | 4 | ABU52951 | Human tra |

ALIGNMENTS

RESULT 1

ADH48764
ID ADH48764 standard; protein; 566 AA.

AC ADH48764;

XX
XX 25-MAR-2004 (first entry)

XX
XX NOV19 protein sequence, SEQ ID 48.

XX Human; NOVX; atherosclerosis; hypertension; obesity; cancer; cytostatic;
KW hypotensive; antiarteriosclerotic; anorectic; gene therapy; NOV19;
KW spinster-like protein; chromosome 17.

XX Homo sapiens.

XX WO200268652-A2.

XX 06-SEP-2002.

XX 26-FEB-2002; 2002WO-US005910.

XX 26-FEB-2001; 2001US-0271646P.

XX 27-FEB-2001; 2001US-0271840P.

XX 28-FEB-2001; 2001US-0272404P.

XX 28-FEB-2001; 2001US-0272405P.

XX 28-FEB-2001; 2001US-0272410P.

XX 28-FEB-2001; 2001US-0272414P.

XX 02-MAR-2001; 2001US-0272878P.

XX 02-MAR-2001; 2001US-0272922P.

XX 02-MAR-2001; 2001US-0273048P.

XX 16-MAR-2001; 2001US-0273300P.

XX 20-MAR-2001; 2001US-0273343P.

XX 20-MAR-2001; 2001US-0278660P.

XX 30-MAR-2001; 2001US-0280039P.

XX 30-MAR-2001; 2001US-0280234P.

XX 02-APR-2001; 2001US-0280818P.

XX 12-APR-2001; 2001US-0283443P.

XX 23-APR-2001; 2001US-0285754P.

XX 24-APR-2001; 2001US-0286096P.

XX 03-MAY-2001; 2001US-0288353P.

XX 17-MAY-2001; 2001US-0291703P.

XX 31-MAY-2001; 2001US-0294834P.

XX 20-JUN-2001; 2001US-0299895P.

XX 21-JUN-2001; 2001US-0299845P.

XX 05-JUL-2001; 2001US-0303242P.

XX 13-AUG-2001; 2001US-0311981P.

stroke, cerebral neoplasms, Pick's disease, Huntington's disease and Parkinson's disease, demyelinating diseases, mental disorders including mood, anxiety, Schizophrenia and seasonal affective disorder, muscle disorder including cardiomyopathy, myocarditis, polymyositis, dermatomyositis, arthromyalgia and asthma and immunological disorders including AIDS, adult respiratory distress syndrome (ARDS), allergies, anemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's syndrome, systemic lupus erythematosus and other diseases including sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, Grave's disease, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, cancers psoriasis and viral, bacterial, fungal, helminthic and protozoal infections. TRICH DNA is useful in gene therapy and in diagnostic purposes

XX SQ Sequence 398 AA;

Query Match 67.7%; Score 1978; DB 4; Length 398;
Best Local Similarity 99.0%; Pred. No. 5.3e-170;
Matches 384; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 152 MVAAPIFGYLDRNRKRVILSCGIFFWSAVTFSSFIPOQYFLLVLSRGLVIGIEASYS 211

Db 1 MVAAPIFGYLDRNRKRVILSCGIFFWSAVTFSSFIPOQYFLLVLSRGLVIGIEASYS 60

Qy 212 TIAPTIIGDLPTKNTKRLTMSVFPALPLGSLGYITGSSVVKQAAGDHWALRVSPVLGM 271

Db 61 TIAPTIIGDLPTKNTKRLTMSVFPALPLGSLGYITGSSVVKQAAGDHWALRVSPVLGM 120

Qy 272 ITGTLIIILVPATKRGHADQGDQKARTSWLRDMKALIRNSYVFSLSATSAVSPATGA 331

Db 121 ITGTLIIILVPATKRGHADQGDQKARTSWLRDMKALIRNSYVFSLSATSAVSPATGA 180

Qy 332 LGMWIPYLHRAQVQKTAETCNPPCGAKDSLIFGAIITCTGFLGVVVTGAGATRWCRK 391

Db 181 LGMWIPYLHRAQVQKTAETCNPPCGAKDSLIFGAIITCTGFLGVVVTGAGATRWCRK 240

Qy 392 TORADPLVCAVGMIGSAIFICLI FVAAKSSIVGAYICIFVGETLLFNFNWAITADILMYV 451

Db 241 TORADPLVCAVGMIGSAIFICLI FVAAKSSIVGAYICIFVGETLLFNFNWAITADILMYV 300

Qy 452 IPTRRATAVALQSTSHLLGDAGSPYLIIGFISDLIROSTKDSPLWEFLSLGYALMLCPFV 511

Db 301 IPTRRATAVALQSTSHLLGDAGSPYLIIGFISDLIROSTKDSPLWEFLSLGYALMLCPFV 360

Qy 512 VVLGGMEFLATLFFVSDRAAEQHLGE 539

Db 361 VVLGGMEFLATLFFVSDRAAEQVQVQ 388

RESULT 3

AD08326

ID AD08326 standard; protein; 312 AA.

XX AC AD08326;

XX AC AD08326;

DT 29-JAN-2004 (first entry)

XX Novel protein (useful for identifying genetic disorders) #481.

DE novel gene; novel protein; tissue marker; molecular weight marker;

XX chromosome marker; genetic disorder.

XX Unidentified.

OS WO2003054152-A2.

XX 03-JUL-2003.

XX 10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

XX 11-DEC-2001; 2001US-0339453P.

PR Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Goodrich RW, Ren P, Zhang J, Zhao QA, Wang J;
XX Ghosh M, Xue AJ, Wehrman T, Wang G, Zhou P, Drmanac RT, Wang Z;
XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.
XX N-PSDB; ADE07415.
XX New polynucleotides, useful for expressing recombinant proteins for
XX analysis, characterization or therapeutic use, or as markers for tissues
XX in which the corresponding protein is preferentially expressed.
XX Claim 20; SEQ ID NO 1392; 1177pp; English.

XX The invention comprises the amino acid and coding sequences of novel
XX proteins. The DNA and protein sequences of the invention are useful as:
XX markers for tissues in which the corresponding protein is preferentially
XX expressed; as molecular weight markers on gels; as chromosome markers or
XX tags; to identify chromosomes or to map related gene positions; and to
XX compare with endogenous DNA sequences in patients to identify potential
XX genetic disorders. The present amino acid sequence represents a protein
XX of the invention.

XX Sequence 312 AA;

Query Match 54.9%; Score 1604; DB 7; Length 312;
Best Local Similarity 99.7%; Pred. No. 2.6e-136;
Matches 310; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MMCLCASAAGAGAEEDAEERRRRRGAQRGAGGSCCGAGGAGGAGVSAAGDEVQTL 60

Db 1 MMCLCASAAGAGAEEDAEERRRRRGAQRGAGGSCCGAGGAGGAGVSAAGDEVQTL 60

Qy 61 SGSVRRAPTPPTPGTCAATAGKGAQPKASLGRGRGAAATLSLGNVNLVDY 120

Db 61 SGSVRRAPTPPTPGTCAATAGKGAQPKASLGRGRGAAATLSLGNVNLVDY 120

Qy 121 TVAGVLLDIQOHFGVXDRGAGLLQSVFICSPMVAAPIFGYLGRFNRKVLSCGIFFWA 180

Db 121 TVAGVLLDIQOHFGVXDRGAGLLQSVFICSPMVAAPIFGYLGRFNRKVLSCGIFFWA 180

Qy 181 VTFSSSFIPOQYFLLVLSRGLVIGIEASYSTIAPTIIIGDLFTKNTKRLTMSVFPFAPL 240

Db 181 VTFSSSFIPOQYFLLVLSRGLVIGIEASYSTIAPTIIIGDLFTKNTKRLTMSVFPFAPL 240

Qy 241 GSGLGIVTSSVKQAAGDHWALRVSPVLGMITGTLILVLPATKRGHADQDGLKART 300

Db 241 GSGLGIVTSSVKQAAGDHWALRVSPVLGMITGTLILVLPATKRGHADQDGLKART 300

Qy 301 SWLRDMKALIR 311

Db 301 SWLRDMKALIR 311

RESULT 4

AAW38697

ID AAW38697 standard; protein; 528 AA.

XX AC AAW38697;

XX AC AAW38697;

DT 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 1842.

DE Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
XX WO200153312-A1.
XX
XX PD 26-JUL-2001.
XX
XX PF 26-DEC-2000; 2000WO-US034263.
XX
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-0058042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSB-) HYSEQ INC.
XX
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX PI Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX
XX DR N-PSDB; AAI57853.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Example 3; SEQ ID NO 1842; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAI38642-AAI42213) with neurotropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 528 AA;
XX
Query Match 44.5%; Score 1298; DB 4; Length 528;
Best Local Similarity 53.4%; Pred. No. 2.4e-108;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;
QY 53 AGDEVQTLGSGVRAPTGP-PGTGTPGCAATAGPGCAQPKPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPGVPPTGCLPGSTGNTGNPKSEPEVPDQEGQLRITGLSPGRSALI 61
QY 105 AAILSLGNVLNLDRTYVAGVLLDIOQHFGVKDRGAGLLQSFTCSFWMVAAPTFGYLGDR 164
DB 62 VAVLCYNLLNMYDRFTVAGVLPDIEQFENIGDSSGLIQTVFTSSYVNLAPVFGYLGDR 121
QY 165 PNRKVLSCGIFFFNSATFSSFTPOQYFWLLVLSRGVIGTGSEASTIAPTIGDLPFK 224
DB 122 YNRKYLKMGCGAFAPFVTLGSSFTPGBFHLLLTGRLGVNGEASTIAPTILADLFVA 181
QY 225 NTRTLMLSVFYFAIPLAGSGLGIYITGSSVYKQAAGDWHWALRVSPVLGMITGTLLILVPAT 284

Db 182 DQSEMLSIYFAIPVSGSLGIYAGSKVQDWDHWRVTPGLGVVAVLLFLVWREP 241
QY 285 KRGHADQIGDQKAR-TSWLRDMKALIRNRSYVFSSLATSAVSFATGALGMWIPLYLHRA 343
Db 242 PRGAVERHSOLPLNPTSMWADLRALRNPSFVLSLSIGFTAVAFVTSGLAWAFALURS 301
QY 344 QVQKTAETC-NSPPCGAKDSLIFGATICTFGFLGVWVTGAGATRCWLKTKORADPLVCV 402
Db 302 RVILGETPPCLPGDSCSSDSLIFGLITCLTGLGVGLGVEISRRLHNSPRADPLVCAT 361
QY 403 GMLGSAIFICLIFVAAKSSIYGVAYICIFVGTLLFSNWAITADILIMYVITTRATAVAL 462
Db 362 GLGSAFPLFLSLACARSGIVATVIFIFIGTLLSMWAIIVADILLYVYVITRSTASAF 421
QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLNEFLSLGVYALMLCPFVVLGGMFFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGLISDLIRRNWPPFLSEFRALQFSLMLCAVFGALGGAFLGT 481
QY 523 ALFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 516
RESULT 5
ABP52155
ID ABP52155 standard; protein; 528 AA.
XX
XX AC ABP52155;
XX
XX DT 14-OCT-2002 (first entry)
XX
XX DE Human 46455 transporter protein SEQ ID NO:5.
XX
XX KW Human; transporter protein; anorectic; antidiabetic; anti-Parkinsonian;
KW neurotropic; neuroprotective; hypotensive; antidepressant; neuroleptic;
KW cardiovascular; immunosuppressive; gene therapy; sugar homeostasis;
KW obesity; diabetes; anorexia; central nervous system disorder; depression;
KW CNS disorder; Parkinson's disease; Alzheimer's disease; hypertension;
KW autonomic function disorder; schizophrenia; learning disorder; amnesia;
KW memory disorder; age-related disorder; cardiovascular disorder;
KW ischaemia reperfusion injury; restenosis; hormonal disorder;
KW hypothyroidism; hyperthyroidism; immune disorder;
KW chronic mucocutaneous candidiasis.
XX
XX OS Homo sapiens.
XX
XX PN WO200255701-A2.
XX
XX PD 18-JUL-2002.
XX
XX PF 17-DEC-2001; 2001WO-US049060.
XX
XX PR 15-DEC-2000; 2000US-0256240P.
XX PR 18-DEC-2000; 2000US-0256588P.
XX PR 21-DEC-2000; 2000US-0258028P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX PI Curtis RAJ;
XX
XX WPI; 2002-590672/63.
XX
XX DR N-PSDB; ABQ74262, ABQ74263.
XX
XX PT New 67076, 67102, 44181, 67084FL and 67084alt nucleic acids and proteins,
XX useful in e.g. treating disorders characterized by insufficient or
XX excessive production of the polypeptides, e.g. cardiovascular or immune
XX disorders.
XX
XX Claim 13; Fig 8; 364pp; English.
XX
XX The present sequence represents the human 46455 protein, which is a
XX member of the transporter family. The transporter family proteins (I) of

the present invention have anorectic, antidiabetic, anti-Parkinsonian, neurotrophic, neuroprotective, hypotensive, antidepressant, neuroleptic, cardiovascular and immunosuppressive activities, and can be used in gene therapy. The human transporter proteins from the present invention are designated 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL and 67084alt. The polynucleotides encoding (I) can be used in screening assays (e.g. chromosome mapping, tissue typing, or in forensic biology), predictive medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical trials, or pharmacogenetics), as surrogate markers, and in methods of treatment (e.g. therapeutic or prophylactic). (I) are useful for treating disorders characterised by insufficient or excessive production of 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt substrates or production of transport 8099, 46455, 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt inhibitors. Disorders associated with 8099 and 46455 include disorders associated with sugar homeostasis such as obesity, diabetes or anorexia. Disorders associated with 54414, 53763, 67076, 67102, 44181, 67084FL or 67084alt are e.g. CNS disorders (e.g. Parkinson's disease or Alzheimer's disease), autonomic function disorders (e.g. hypertension, depression or schizophrenia), or learning or memory disorders (e.g. amnesia or age-related disorders), cardiovascular disorders (e.g. ischaemia reperfusion injury or restenosis), hormonal disorders (e.g. hypothyroidism or hyperthyroidism), or immune disorders (e.g. chronic mucocutaneous candidiasis)

Sequence 528 AA:

[illegible]

RESULT 6

RESUL 6
ADD37478

ADD37478 standard; protein: 528 AA.

XX

AC ADD37478;

| | | |
|----|-----------------------------------|---------------|
| DT | 15-JAN-2004 | (first entry) |
| XX | | |
| DE | Human transporter | 46455. |
| XX | | |
| KW | Human; transporter; cytochrome | |
| KW | gene therapy; PGC-1 associated | |
| KW | diabetes. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | US2003143675-A1. | |
| XX | | |
| PD | 31-JUL-2003. | |
| XX | | |
| PF | 22-MAY-2002; 2002US-00154419. | |
| XX | | |
| PR | 12-MAY-2000; 2000US-0204211P. | |
| PR | 29-JUN-2000; 2000US-0215376P. | |
| PR | 31-JUL-2000; 2000US-0221769P. | |
| PR | 19-SEP-2000; 2000US-0237907P. | |
| PR | 25-SEP-2000; 2000US-0235107P. | |
| PR | 05-OCT-2000; 2000US-0239336P. | |
| PR | 14-NOV-2000; 2000US-0248364P. | |
| PR | 15-NOV-2000; 2000US-0248819P. | |
| PR | 15-DEC-2000; 2000US-0256240P. | |
| PR | 18-DEC-2000; 2000US-0255588P. | |
| PR | 21-DEC-2000; 2000US-0258028P. | |
| PR | 22-JAN-2001; 2001US-0263169P. | |
| PR | 14-MAY-2001; 2001US-0085811P. | |
| PR | 29-JUN-2001; 2001US-00895811. | |
| PR | 31-JUL-2001; 2001US-00919781. | |
| PR | 19-SEP-2001; 2001US-0095766P. | |
| PR | 25-SEP-2001; 2001US-0096429S. | |
| PR | 05-OCT-2001; 2001US-0097272A. | |
| PR | 14-NOV-2001; 2001US-00902769. | |
| PR | 17-DEC-2001; 2001US-00024623. | |
| PR | 22-JAN-2002; 2002US-0005502S. | |
| XX | | |
| PA | (WILL-) MILLENNIUM PHARM INC. | |
| XX | | |
| PI | Curtis RAJ, Glucksmann MA, et al. | |
| PI | WPI; 2003-851783/79. | |
| DR | N-PSDB; ADD37477, ADD37479. | |
| DR | | |
| PT | New isolated nucleic acid, use | |
| PT | treating PGC-1 associated disor | |
| PT | or diabetes. | |
| XX | | |
| PS | Claim 11; SEQ ID NO 55; 663b | |

Query Match 44.5%: Score 1298: DB 7: Length 528:

Best Local Similarity 53.4%; Pred. No. 2.4e-108;

| | | | | | | | | | |
|---------|------|--------------|-----|------------|------|--------|-----|------|----|
| Matches | 275; | Conservative | 62; | Mismatches | 162; | Indels | 16; | Gaps | 7; |
|---------|------|--------------|-----|------------|------|--------|-----|------|----|

```
QY 53 AGDEVQTLSGSVRRAPTGP-PCTGPTGCAATAKAGPQAQPKPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGVPCTGPGTGNPKSEEPVQDGLQRTIGLSPGRSALI 61
QY 105 AAILSLGNVNLDRYTVAGVLLDQOHFGVDRGAGLQSVFICSFMAAPFVGLGR 164
DB 62 VAVLCYINLLNMDRFTVAGVLPDIEQFNIGDSSGLQTVFISYMWLAFVFGVGLGR 121
QY 165 FNRKVLSCGIPFWSAVTFSSFPQYFWLLVLSRGLVIGGEASYSTIAPTIGDLFTK 224
DB 122 YNRKYLSCGIAFWLSVLTGSSFIPEHFVLLLTGLVGVGEASYSTIAPTIALDFVA 181
QY 225 NTRTLMVSYFAIPGLSGGLVITGSSVKQAAGDHWALRVSPVLGMITGTLILVPAT 284
DB 182 DQSRMLSFYFAIPVGSGLGYIAGSKVDMAGDHWALRVTPGLGVAVALLFLVVRP 241
QY 285 KEGHADQLDQLKAR-TSWLRDMKALIRNRSVVFSSLATSAVSFATGALGMWIPLYLHRA 343
DB 242 PRGAVERHSDLPPLNPTSWADRLARNPFSVLSLGFATAVFTGSLALWAPALLRS 301
QY 344 QVVQKTAETC-NSPPCGAKDSLIFGAITCTFTGFLGVVTVGAGATRCRLKTRQADPLVCV 402
DB 302 RVVLGETPCLPGDCSSSDSLIFGLITCLTGLVGLGVETISRRLRHSNPRADPLVCAT 361
QY 403 GMLGSAIFCLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTERTAVAL 462
DB 362 GLLGSAFFLFLSLACARGSIATVPIFIFIGETLLSNMNAIVADILLYVVIPTRRSTAEAF 421
QY 463 QSFTHLLGDAGSPYLIGTISDLIRQSTKDSPLWBFSLGVALMLCPFFVVVLGSGMFFLAT 522
DB 422 QIVLSHLLGDAGSPYLIGLISDLRLRNWPPSFLSEPRALQPSLMCAFVGLGGAFLGT 481
QY 523 ALFFVSDRABAROHGSG--RRAGVR-----VHQRG 551
DB 482 AIFIADRRRAQLHVQGLLHEAGSTDTRIWWPQRG 516

RESULT 7
AD127957
ID AD127957 standard; protein; 528 AA.
XX
AC AD127957;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human 46455 protein.
XX
KW Human; 46455; ion channel family; ICF; cancer; leukaemia;
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; epilepsy;
KW hepatic disorder; cardiovascular disorder; cytostatic; neuroprotective;
KW nootropic; antiparkinsonian; hepatotropic; cardiovascular.
XX
OS Homo sapiens.
XX
PN US2003165891-A1.
XX
PD 04-SEP-2003.
XX
PF 15-MAY-2002; 2002US-00146733.
XX
PR 29-FEB-2000; 2000US-00515520.
PR 29-FEB-2000; 2000US-0185338P.
PR 03-MAR-2000; 2000US-00518866.
PR 07-APR-2000; 2000US-0195734P.
PR 11-APR-2000; 2000US-0195993P.
PR 26-APR-2000; 2000US-0199799P.
PR 19-SEP-2000; 2000US-0233537P.
PR 25-SEP-2000; 2000US-0235018P.
PR 25-SEP-2000; 2000US-0235059P.
PR 18-DEC-2000; 2000US-0256240P.
PR 21-DEC-2000; 2000US-0256588P.
PR 28-FEB-2001; 2001US-00796720.
```

```
PR 06-APR-2001; 2001US-00828035.
PR 11-APR-2001; 2001US-00833081.
PR 25-APR-2001; 2001US-00843128.
PR 19-SEP-2001; 2001US-00957683.
PR 25-SEP-2001; 2001US-00964252.
PR 25-SEP-2001; 2001US-00964256.
PR 17-DEC-2001; 2001US-00024623.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtis RAJ, Glucksmann MA, Silos-Santiago I;
XX
XX WPI: 2004-069000/07.
XX N-PSDB; AD127956, AD127958.
XX
PT TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9,
PT alpha2delta-4, 54414, or 53763 nucleic acids and proteins, useful for
PT preventing, diagnosing and treating e.g. cancers, Alzheimer's disease and
PT cardiovascular disorders.
XX
PS Example 1; SEQ ID NO 50; 638pp; English.
XX
CC The invention relates to TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611,
CC IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and 53763 ion channel
CC family (ICF) nucleic acids and proteins. The TWIK-6, TWIK-7, IC23927,
CC TWIK-8, IC47611, IC47615, HNMDA-1, TWIK-9, alpha-2delta-4, 54414 and
CC 53763 ICF nucleic acids and proteins may be used for preventing,
CC diagnosing and treating ICF-related diseases. The sequences may be used
CC to treat disorders associated with decreased expression by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC ICF proteins by expressing inactive proteins or to supplement the
CC patients own production of ICF proteins. The proteins may also be used as
CC antigens in the production of antibodies against ICF proteins and in
CC assays to identify modulators of ICF protein expression and activity. The
CC anti-ICF protein antibodies, agonists and antagonists may be used to
CC regulate ICF protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of ICF proteins in
CC samples (e.g. by immunoassay). The nucleic acids and proteins may be used
CC to prevent, diagnose and treat a wide variety of disorders, e.g. cancers
CC and leukaemia, Alzheimer's disease, Parkinson's disease, multiple
CC sclerosis, epilepsy, hepatic disorders and cardiovascular disorders. This
CC sequence represents the human 46455 protein of the invention. Note: The
CC sequence data for this patent is also available in electronic format from
CC USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 528 AA;
Query Match 44.5%; Score 1298; DB 8; Length 528;
Best Local Similarity 53.4%; Pred. No. 2.4e-108;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;
QY 53 AGDEVQTLSGSVRRAPTGP-PCTGPTGCAATAKAGPQAQPKPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGVPCTGPGTGNPKSEEPVQDGLQRTIGLSPGRSALI 61
QY 105 AAILSLGNVNLDRYTVAGVLLDQOHFGVDRGAGLQSVFICSFMAAPFVGLGR 164
DB 62 VAVLCYINLLNMDRFTVAGVLPDIEQFNIGDSSGLQTVFISYMWLAFVFGVGLGR 121
QY 165 FNRKVLSCGIPFWSAVTFSSFPQYFWLLVLSRGLVIGGEASYSTIAPTIGDLFTK 224
DB 122 YNRKYLSCGIAFWLSVLTGSSFIPEHFVLLLTGLVGVGEASYSTIAPTIALDFVA 181
QY 225 NTRTLMVSYFAIPGLSGGLVITGSSVKQAAGDHWALRVSPVLGMITGTLILVPAT 284
DB 182 DQSRMLSFYFAIPVGSGLGYIAGSKVDMAGDHWALRVTPGLGVAVALLFLVVRP 241
QY 285 KEGHADQLDQLKAR-TSWLRDMKALIRNRSVVFSSLATSAVSFATGALGMWIPLYLHRA 343
DB 242 PRGAVERHSDLPPLNPTSWADRLARNPFSVLSLGFATAVFTGSLALWAPALLRS 301
QY 344 QVVQKTAETC-NSPPCGAKDSLIFGAITCTFTGFLGVVTVGAGATRCRLKTRQADPLVCV 402
```

Db 302 RVLGTPPCLPDSCSSDLSIFGLITCLTGLVGLGVGEISRLRHSNPRADPLVCAT 361
 QY 403 GMLGSAIFICLI FVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462
 Db 362 GLLGSAPFLSLACARGSI VATYIFIFIGETLLSMNWAIVADILLYVVIPTTRSTAEAF 421
 QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSKDPSPLWEFLSLGVALMCPVVVLGGMFFLAT 522
 Db 422 QIVLSHLLGDAGSPYLIGLISDRUNWPPSFLSEFRALQPSLMCAFPVGLGAAFLGT 481
 QY 523 ALFFVSDRRAEQLHGE--RRAGVR-----VVHQRG 551
 Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIVVVPPQRG 516

RESULT 8
 ADL61327
 ID ADL61327 standard; protein; 528 AA.
 XX
 AC ADL61327;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Human protein tyrosine kinase biomarker spinster-like protein.
 XX
 KW predictor set; protein tyrosine kinase; cytostatic; antiangiogenic;
 KW vasotrophic; vulnery; pharmacogenomic; drug sensitivity; breast cancer;
 KW hypervascular disease; angiogenesis; wound healing scar; human;
 KW biomarker; spinster-like.
 XX
 OS Homo sapiens.
 XX
 PN WO2004020583-A2.
 XX
 PD 11-MAR-2004.
 XX
 PF 26-AUG-2003; 2003WO-US026491.
 XX
 PR 27-AUG-2002; 2002US-0406385P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Huang F, Han X, Reeves KA, Amler L, Fairchild CR, Lee FY;
 PI Shaw P;
 XX
 DR WPI; 2004-239171/22.
 DR N-PSDB; ADL61203.
 XX
 XX New predictor sets with a plurality of polynucleotides and/or
 PT polypeptides whose expression pattern predicts cell response to a
 PT compound that modulates protein tyrosine kinase activity, useful in
 PT treating breast cancer.
 XX
 PS Claim 9; SEQ ID NO 251; 649pp; English.
 XX

The invention relates to a novel predictor set comprising a plurality of polynucleotides and/or polypeptides whose expression pattern is predictive of the response of cells to treatment with a compound that modulates protein tyrosine kinase activity or members of the protein tyrosine kinase pathway. The molecules of the invention demonstrate cytostatic, antiangiogenic, vasotrophic and vulnery activities and may be useful in the field of pharmacogenomics, in particular for determining drug sensitivity and in treating breast cancer, hypervascular diseases, angiogenesis and scars in wound healing. The current sequence is that of a human protein tyrosine kinase biomarker protein of the invention.

XX
 SQ Sequence 528 AA;
 Query Match 44.5%; Score 1298; DB 8; Length 528;
 Best Local Similarity 53.4%; Pred. No. 2.4e-108;
 Matches 275; Conservative -62; Mismatches 162; Indels 16; Gaps 7;
 QY 53 AGDEVQTLGSGVRRAPTGP-PGTGTPGCAATAKGAGQAQPKFASLGR-----GRGA-A 104,

Db 2 AGSDTAPFLSOADDDPDGPPGTFGLPGSTGNPKSEPEVPDQEGRLQITGLSPERSALI 61
 QY 105 AAILSLGNLVNLDYRTVAVGLLDIOQHFGVKDRAGLQSVFICSFVAAAPIFGYLGDR 164
 Db 62 VAVLCYINLLNNDYRTVAVGLPDIEQFNIGDSSGLIQTFFISSYVNLAPVFGYLGDR 121
 QY 165 FNKRVLUSCGIFPWSAVTFSSSFIPOQYFWLLVLSRGLVGIGEASYSTIAPTIIGDLPFK 224
 Db 122 YNRKYLKCGGTAFWSLVTLGSSSFIPEGHFHLLLLITRGLVGVEASYSTIAPTLLADLFVA 181
 QY 225 NTRTLMLSVFPFAIPGLSGLYITGSSVKQAAGDMHWALRVSPVLGMITGTLLILLYPAT 284
 Db 182 DQSRMLSIIFYFALPVGSLGYTAGSKYKQWAGDMHWALRVTPGLGVVAVLLFLVWREP 241
 QY 285 KRGHADQLGQDKAR-TSWLRDMKALIRNRSYVFFSLATSAVSPATGALGMWIPLYLHRA 343
 Db 242 PRGAVEHSDLPPLNPISWADIRALARNPSFVLSSLGFTAVAFVTSGLMWAFAFLRS 301
 QY 344 QVVKQTAETC-NSPPCGAKDSLIFGATCTFTGLGVVTVGAGATRWCKLKTORADPLVCV 402
 Db 302 RVLGTPPCLPDSCSSDLSIFGLITCLTGLVGLGVGEISRLRHSNPRADPLVCAT 361
 QY 403 GMLGSAIFICLI FVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462
 Db 362 GLLGSAPFLSLACARGSI VATYIFIFIGETLLSMNWAIVADILLYVVIPTTRSTAEAF 421
 QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSKDPSPLWEFLSLGVALMCPVVVLGGMFFLAT 522
 Db 422 QIVLSHLLGDAGSPYLIGLISDRUNWPPSFLSEFRALQPSLMCAFPVGLGAAFLGT 481
 QY 523 ALFFVSDRRAEQLHGE--RRAGVR-----VVHQRG 551
 Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIVVVPPQRG 516

RESULT 9
 ADO57285
 ID ADO57285 standard; protein; 507 AA.
 XX
 AC ADO57285;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Kidney development associated protein seqid 52.
 XX
 KW nephrotropic; cardiant; antiinfertility; cerebroprotective;
 KW neuroprotective; muscular; cytostatic; osteopathic; gene therapy;
 KW kidney development; kidney disorder; developmental disorder;
 KW circulatory disorder; hearing disorder; heart defect; infertility;
 KW stroke; mental retardation; muscle defect; proliferative disorder;
 KW bone defect; bone disorder; zebrafish.
 XX
 OS Danio rerio.
 XX
 PN US2004068763-A1.
 XX
 PD 08-APR-2004.
 XX
 PF 28-MAR-2003; 2003US-00403571.
 XX
 PR 29-MAR-2002; 2002US-0368760P.
 XX
 PA (HOPK/) HOPKINS N.
 PA (GOLL/) GOLLING G.
 PA (AMST/) AMSTERDAM A.
 PA (SUNZ/) SUN Z.
 XX
 PI Hopkins N, Golling G, Amsterdam A, Sun Z;
 XX
 DR WPI; 2004-304692/28.
 DR N-PSDB; ADO57284.
 XX

PT New 459 nucleic acids and encoded polypeptides, useful for diagnosing,
PT treating or preventing a kidney disorder in an organism, or in screening
PT for compounds that modulate the development of an organism.

XX Disclosure; SEQ ID NO 52; 347pp; English.

XX The invention describes an isolated nucleic acid molecule (I) comprising
CC a sequence having at least 75% sequence identity to the 459 nucleic acid
CC sequence of 2808 base pairs (SEQ ID NO: 59) given in the specification,
CC over at least 600 contiguous base pairs, where the nucleic acid functions
CC in kidney development. (I) is useful for treating or preventing a kidney
CC disorder in an organism, where the nucleic acid elicits an alteration in
CC expression of a 459 nucleic acid sequence in the organism and
CC subsequently treats or prevents a kidney disorder. The nucleic acid may
CC also be used in diagnosing, preventing and treating a variety of
CC mammalian diseases and developmental disorders (e.g. circulatory
CC disorders, hearing disorders, heart defect, infertility, stroke, mental
CC retardation, muscle defects, proliferative disorders, or bone defects or
CC disorders) as well as in screening for compounds that modulate the
CC development of an organism as a whole or of specific tissues or organs
CC within that organism. This is the amino acid sequence of a kidney
CC development associated protein.

XX Sequence 507 AA;

Query Match 43.0%; Score 1256.5; DB 8; Length 507;

Best Local Similarity 56.4%; Pred. No. 1.3e-104;

Matches 251; Conservative 66; Mismatches 125; Indels 3; Gaps 2;

QY 107 ILSGLNVLNLDYRTVAGVLLDIOQHFGVKDRGAGLLQSVFCSFMAPIFGYLGDRFN 166

DB 54 VLYCYNLLNMDRFTVAGVLPDIEHFFGIGDGTSGLLQTVFICSYMFAPLFGYLGDRYN 113

QY 167 RKVILSCGIFFSAVTFSSFPQQYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTKNT 226

DB 114 RKLIMCVGIFFSVWTLASSFIGKHDFWALLTRGLVGIGEASYSTIAPTIIADLTVKEK 173

QY 227 RTMLSVFVFAIPGLSGLOYITGSSVKQAAGDWHALRVSPVLGMITGTLLIILPATKR 286

DB 174 RTNMLSIIFYFAIPVSGMGYIVGSKVDITVAKDWHALRVTPGLGLLAVFLLMLVVOEPR 233

QY 287 GHADQLGDQKARTWLRDMKALINRSVVFSSLATSVSFATGALGMWIPLYLHRAQV 346

DB 234 GAIEAHPETHLRTSWLMDKALCRNPGFILSTFGFVAFTGSLAWAPLFRAGVF 293

QY 347 QKTAETCNSPPCGAKDSLIFGAITCTGFLGVVGTGATWCRCLKTQRADPLVCVGMUG 406

DB 294 TGVKQPCFKAPCDDSDSLIFGAITVTVTGILGVASGVQASKLQRTTRPRADPLVCAAGLL 353

QY 407 SAIFICLIFFAAKSSIVGAYICIFVGETLLFSNWAITADILMVVITPTTRATAVALQSFT 466

DB 354 AAPFLYLSIMFAQASTVATVYFIFLGETFLSNWAIVADILYVVIPTRRTAEAPQIVL 413

QY 467 SHLLGAGSPYILGIFTSIDLIROSTKDSPLWEFLSLGYALMLCPFVVVLGCMFFLATLRF 526

DB 414 SHLLGDAISPYILGVVSDSIKES--NSYMWFRSLQMSLLCSFVPVAGGAFFLATAVFI 471

QY 527 VSDRAEAQHLGERRAGVVRVHQRG 551

DB 472 EKORDLAENYVPSDDAPI-VVPRSG 495

RESULT 10

AAM40483

ID AAM40483 standard; protein; 581 AA.

XX AAM40483;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 5414.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSEQ-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI59639.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

XX as central nervous system injuries.

XX Example 2; SEQ ID NO 5414; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic, and the
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX Sequence 581 AA;

Query Match 42.8%; Score 1250; DB 4; Length 581;

Best Local Similarity 50.9%; Pred. No. 6e-104;

Matches 268; Conservative 63; Mismatches 152; Indels 44; Gaps 7;

QY 8 SAAAGCAEEEADEARRRRRRRGAQAGGSG--CCGARGAGAGVSAAGDEVOTLGSVRR 66

DB 85 TGAAGAPRDHGRVHRHRPPSSARRMTRTGTQCLAPRG----- 123

QY 67 APTGPPGT--PGTP-----GCAATAKGPAGQOPKPSASLGRGRGAAATLSLGNVLNYLD 118

DB 124 --QGPRGTRSPRSPRSTRTRRGCSAS-----PACLPCSRALIVAVLCINLLNYMD 171

QY 119 RYTVAGVLLDIOQHFGVKDRGAGLLQSVFICSFMAAPIFGYLGRFNRKVLSCCFFWF 178

DB 172 RFTVAGVLPDIEQFFNIGDSSGLLQTVFISSYMWLAPVFGYLGDRYKRYLMCGGIAFW 231

QY 179 SAVTFSSSFIPOQYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTKNTTLMLSVFPFI 238

XX 18-FEB-2000; 2000WO-US004340.
 PF 19-FEB-1999; 99US-0120680P.
 XX 23-APR-1999; 99US-00298733.
 PR 17-AUG-1999; 99US-0149639P.
 PR 23-SEP-1999; 99US-0155686P.
 PR 01-OCT-1999; 99US-0157247P.
 PR 29-NOV-1999; 99US-0167822P.
 PR 29-NOV-1999; 99US-0167823P.
 PR 15-FEB-2000; 2000US-0182711P.
 XX (ALPH-) ALPHAGENE INC.
 XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 XX WPI; 2000-549267/50.
 DR N-PSDB; AAA93110.
 XX New secreted proteins and polynucleotides encoding them, which are
 PT derived from Homosapiens, useful for therapy, diagnosis, and research, as
 PT well as nutritional sources or supplements.
 XX Claim 29; Page 254; 309pp; English.
 XX The present sequence is the sequence of a human secreted protein. Its
 CC cDNA was isolated from an adult pancreas cDNA library. The proteins and
 CC coding sequences of the invention can be used in the isolation of similar
 CC genes and proteins, in the elucidation of their function in vivo, and to
 CC treat a number of conditions. It is possible that they may have uses as
 CC nutritional supplements, as cytokine or cell proliferation factors, in
 CC immune modulation, where they may be used to treat immune and autoimmune
 CC diseases, as haematopoiesis regulators (treating myeloid or lymphoid cell
 CC deficiencies), in the promotion of tissue growth, they may have chemokine
 CC or chemotactic activity, haemostatic or thrombolytic activity, or anti-
 CC inflammatory activity
 XX Sequence 244 AA;
 SQ
 Query Match 40.3%; Score 1176; DB 3; Length 244;
 Best Local Similarity 97.9%; Pred. No. 9.8e-98;
 Matches 229; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 306 MKALINRSVPSLSATSAVSPATGALGWIPYLHRAQVVKTAETCNPPCGAKDSLII 365
 DB 1 MKALINRSVPSLSATSAVSPATGALGWIPYLHRAQVVKTAETCNPPCGAKDSLII 60
 QY 366 FGAICTCTGFLGVVGTGAGATRCRLKTQRADPLVCAGVGLGSAIFCLIFVAAKSIVGA 425
 DB 61 FGAICTCTGFLGVVGTGAGATRCRLKTQRADPLVCAGVGLGSAIFCLIFVAAKSIVGA 120
 QY 426 YICIFVGETLLFSNWAITADILMYVPIPTRRATAVALQSFTSHLLGDAGSPYLIGFISDL 485
 DB 121 YICIFVGETLLFSNWAITADILMYVPIPTRRATAVALQSFTSHLLGDAGSPYLIGFISDL 180
 QY 486 IROSTKDSPLWEFLSLGYALMLCPFFVVLGGMFFLATLFFVSDRARAEQHLGE 539
 DB 181 IROSTKDSPLWEFLSLGYALMLCPFFVVLGGMFFLATLFFVSDRARAEQVQVQ 234
 RESULT 13
 AAE06616
 ID AAE06616 standard; protein; 476 AA.
 XX AAE06616;
 AC AAE06616;
 XX 25-SEP-2001 (first entry)
 XX Human protein having hydrophobic domain, HP10735.
 XX Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 XX multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;

KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 XX contraceptive; antiinfertility; antiinflammatory.
 OS Homo sapiens.
 XX WO200149728-A2.
 PN 12-JUL-2001.
 PD 28-DEC-2000; 2000WO-JP009359.
 XX 06-JAN-2000; 2000JP-00000585.
 PR 06-JAN-2000; 2000JP-00000588.
 PR 11-JAN-2000; 2000JP-00002299.
 PR 03-FEB-2000; 2000JP-00026862.
 PR 03-MAR-2000; 2000JP-00058367.
 XX (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA Kato S, Kimura T;
 PI WPI; 2001-418355/44.
 DR N-PSDB; AAD12611.
 XX Human proteins with hydrophobic domains and the nucleic acids encoding
 PT them, useful for preventing diagnosing and treating e.g. cancer,
 PT Alzheimer's and inflammation.
 XX Claim 1; Page 150-151; 563pp; English.
 XX The present sequence is human protein with hydrophobic domain, HP10735.
 CC The polynucleotide and polypeptide of the invention may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate polypeptide expression. The polynucleotides may be used to
 CC produce the polypeptide, by inserting the nucleic acids into a host cell
 CC and culturing the cell to express the protein. The polynucleotides and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays and also used in gene therapy. The polypeptides may also be used
 CC as antigens in the production of antibodies and in assays to identify
 CC modulators of polypeptide expression and activity. The polypeptides and
 CC nucleic acids may be used as nutritional supplements, to modulate
 CC cytokine and cell proliferation activity, to modulate immune stimulation
 CC or suppression (e.g. for the treatment of microbial infections and
 CC autoimmune disorders such as multiple sclerosis, rheumatoid arthritis and
 CC insulin-dependent diabetes), to modulate haematopoiesis, to modulate
 CC tissue growth activity (e.g. for the treatment of Parkinson's disease,
 CC Huntington's disease and Alzheimer's disease), to modulate activin and
 CC inhibin activity (e.g. for controlling fertility), to modulate
 CC chemotactic and chemokinetic activity, to modulate haemostatic and
 CC thrombolytic activity, to modulate receptor ligand activity, to modulate
 CC inflammation and to inhibit tumour growth
 XX Sequence 476 AA;
 SQ
 Query Match 40.0%; Score 1167; DB 4; Length 476;
 Best Local Similarity 49.4%; Pred. No. 1.5e-96;
 Matches 254; Conservative 53; Mismatches 141; Indels 66; Gaps 7;
 QY 53 AGDEVOTLGSVRRAPTGP-PGTPTPGCAATAKGPQAQPKPSLGR-----GRGA-A 104
 DB 2 AGSDTAPFLSQADDDPDGPFVGTGTLPGSTGNPKSEEPVDPQEQLOITGLSPERSALI 61
 QY 105 AAILSLGNVLNLDRTYVAGVLLDIQHQHFGVKDRGAGLLQSVFICSFVWAPIFGYLGR 164
 DB 62 VAVLCYINLLNMDRFTVAGVLPDIEQPFNIGDSSGLIQTVFISYVWVAPVFGYLGR 121
 QY 165 FNRKVLSCGIFFWSAVTFSSSFPQQYFWLLVLSRGLVGEASYSYTIPTIGDLTK 224
 DB 122 YNRKYLKMGCGIAFWLSVTLGSSFIPGEHFHLLLLTRGLVGVGEASYSYTIPTLADLFA 181

QY 524 LFFVSDRAAEQHLGE--RRAGVR-----VHQRG 551
Db 431 IFIEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 464

QY 450 VVIPTRRATAVALOSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSIGYALMLCP 509
Db 301 VVIPTRRSTAEAFQIVLSHLLGDAGSPYLIGLISDLRLARNWPPSFLSEFRALQFSMLCA 360
QY 510 FVVVLGGMFFLATATLFFVSDRAAEQHLGE--RRAGVR-----VHQRG 551
Db 361 FVGALGGAAFLGTAFIEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 408

Search completed: October 23, 2004, 13:12:09
Job time : 93 secs

RESULT 15

ADL08448
ID ADL08448 standard; protein; 420 AA.

XX AC ADL08448;

XX DT 06-MAY-2004 (first entry)

XX DE Human cancer suppressor protein PP2030.

XX KW cancer suppressor; cancer.

XX OS Homo sapiens.

XX PN CN1403480-A.

XX PD 19-MAR-2003.

XX PF 12-SEP-2001; 2001CN-00126728.

XX PR 12-SEP-2001; 2001CN-00126728.

XX PA (SHAN-) SHANGHAI XINSHIJI GENE TECHN DEV CO LTD.

XX PI Gu J, Yang S;

XX DR WPI; 2003-494229/47.

XX DR N-PSDB; ADL08449.

XX PT Human protein with function of suppressing cancer cell growth and its coding sequence.

XX PS Claim 1; SEQ ID NO 2; 42pp; Chinese.

XX CC The invention discloses one kind of human protein with cancer suppressing function, polynucleotides encoding the polypeptide and the recombinant process of producing the polypeptide. The polypeptide can be used in treating various diseases, such as cancer. The present invention also discloses the agonist resisting the polypeptide and its treatment effect. This sequence represents one of the cancer suppressing proteins.

XX SQ Sequence 420 AA;

Query Match 38.2%; Score 1115; DB 7; Length 420;
Best Local Similarity 56.6%; Pred. No. 6.4e-92;

Matches 231; Conservative 50; Mismatches 119; Indels 8; Gaps 4;

QY 152 MVAAPIFGYLGRFRNKRKVLSCGIFPWSAVTFSSSFIPOQYFWLLVLSRGLVIGEAYS 211
Db 1 MVLAPVFGYLDYRNKRYKMGCIAPFWSLVTLGSSFIPEGHFHFWLLLTGRLVGVEAYS 60

QY 212 TIAPTILIGLFTKNTFTMLSVFYFAIPLGSLGXYITGSSVKQAGDWHALRVSPVLGM 271
Db 61 TIAPTILADLFVADQKRMLSIFFYFAIPVSGSLGIAGSKVDMAGDWHALRVTPGLGV 120

QY 272 ITGTILILVPATKRGHADQLGDKAR-TSWLRDMKALIRNSVYFSSLATSAVSFATG 330
Db 121 VAVLLFLVVRPPRGAVRHSDDLPLNFTSNWADLRALARNPSFVLSLGFATAVFTG 180

QY 331 ALGMWIPLYLHRAQVVKQTAETC-NSPPCGAKDSLIFGAITCTGTGFLGVVTCAGATMCR 389
Db 181 SLALWAPAFLLRSRVVLGETPPCLPGDCSSDSLIIFGLITCLTGLVGLGVEISRRLR 240

QY 390 LKTRADPLVCAGVLGSAIFCLIFVAKSSIIVGAYICIFVGETLLFSNWAITADILMY 449
Db 241 HSNPRADPLVCATGLGSAFFLFLACARGSIATYIFIFIGETLLSNWAIADILLY 300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 13:10:39 ; Search time 27 Seconds
(without alignments)

1390.222 Million cell updates/sec

Title: US-10-085-198-48

Perfect score: 2320

Sequence: 1 MMCLCSAAGGAEEAD.....VHQRGPGTALAHVVVGAS 566

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 779 | 26.7 | 299 | 4 | US-09-149-476-396 |
| 2 | 672 | 23.0 | 424 | 4 | US-09-149-476-555 |
| 3 | 396 | 13.6 | 136 | 4 | US-09-149-476-554 |
| 4 | 250.5 | 8.6 | 458 | 4 | US-09-328-352-7272 |
| 5 | 206.5 | 7.1 | 352 | 4 | US-09-252-991A-30568 |
| 6 | 206.5 | 7.1 | 430 | 4 | US-09-489-039A-13997 |
| 7 | 202.5 | 6.9 | 488 | 4 | US-10-162-012-46 |
| 8 | 197.5 | 6.8 | 408 | 4 | US-09-328-352-5768 |
| 9 | 189.5 | 6.5 | 443 | 4 | US-09-489-039A-9335 |
| 10 | 188 | 6.4 | 506 | 4 | US-09-252-991A-17560 |
| 11 | 186 | 6.4 | 463 | 4 | US-09-489-039A-10827 |
| 12 | 183.5 | 6.3 | 503 | 4 | US-09-489-039A-10722 |
| 13 | 182 | 6.2 | 444 | 4 | US-09-492-709A-258 |
| 14 | 180 | 6.2 | 442 | 4 | US-09-489-039A-12384 |
| 15 | 176.5 | 6.0 | 425 | 4 | US-09-711-164-428 |
| 16 | 173.5 | 5.9 | 423 | 4 | US-09-328-352-6097 |
| 17 | 173 | 5.9 | 455 | 4 | US-09-489-039A-9942 |
| 18 | 173 | 5.9 | 433 | 4 | US-09-489-039A-11835 |
| 19 | 168 | 5.8 | 446 | 4 | US-09-489-039A-7920 |
| 20 | 168 | 5.8 | 446 | 4 | US-09-328-352-7326 |
| 21 | 167 | 5.7 | 463 | 4 | US-09-252-991A-29935 |
| 22 | 166.5 | 5.7 | 443 | 4 | US-09-328-352-7567 |
| 23 | 166.5 | 5.7 | 460 | 4 | US-09-252-991A-28387 |
| 24 | 165.5 | 5.7 | 500 | 4 | US-09-252-991A-25053 |
| 25 | 165.5 | 5.7 | 502 | 4 | US-09-489-039A-13185 |
| 26 | 165 | 5.7 | 415 | 4 | US-09-489-039A-10457 |
| 27 | 163.5 | 5.6 | 443 | 4 | US-09-489-039A-9381 |

Sequence 6461, Ap
Sequence 9663, Ap
Sequence 5305, Ap
Sequence 4726, Ap
Sequence 10203, A
Sequence 7076, Ap
Sequence 32216, A
Sequence 25836, A
Sequence 17, Appl
Sequence 2084, Ap
Sequence 8, Appl
Sequence 5594, Ap
Sequence 6006, Ap
Sequence 11902, A
Sequence 10630, A
Sequence 3422, Ap
Sequence 5706, Ap

28 163 5.6 418 4 US-09-328-352-6461
29 162.5 5.6 460 4 US-09-489-039A-9663
30 162 5.5 401 4 US-09-543-681A-5305
31 162 5.5 449 4 US-09-328-352-4726
32 161.5 5.5 476 4 US-09-489-039A-10203
33 161 5.5 421 4 US-09-543-681A-7076
34 160.5 5.5 537 2 US-08-647-397-2
35 160.5 5.5 559 4 US-09-252-991A-32216
36 160 5.5 664 4 US-09-252-991A-25836
37 158.5 5.4 484 4 US-09-266-965-17
38 158 5.4 379 4 US-09-710-279-2084
39 158 5.4 848 4 US-09-575-081B-8
40 157.5 5.4 393 3 US-09-134-001C-5594
41 157.5 5.4 533 4 US-09-107-532A-6006
42 157 5.4 514 4 US-09-489-039A-11902
43 156.5 5.4 459 4 US-09-489-039A-10630
44 156 5.3 321 4 US-09-134-000C-3422
45 155.5 5.3 475 4 US-09-328-352-5706

ALIGNMENTS

RESULT 1

US-09-149-476-396
; Sequence 396, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23


```
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match      26.78; Score 779; DB 4; Length 299;
Best Local Similarity 56.18; Pred. No. 1.5e-61;
Matches 162; Conservative 34; Mismatches 91; Indels 2; Gaps 2;

QY 230 MLSVFYFAIPGLSGLYITSSYKQAGDHWALRVSPVLGMITGTILILVATKRGHA 289
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MLSIFFAIPVSGGLGYAGSKVDAGDHWALRVTPGLGVAVLLFLVVRPPRGAV 60
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 290 DQGDQLKAR-TSWLRDMKALINRNSYVFSLSATSAVSFATGALGMWIMPLYLHRAOVQVK 348
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ERHSDLPPLNPTSWADRLARALNPSPVLSLGTAVAFVTSGLAWAPAFILRSRVLG 120
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 349 TATC-NSPPCGAKDSLIFGAITCTGFLGVTVTGAGATRCWLKTRADPLVCAVGMG 407
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 ETPECLPGSCSSDSLIFGLITCLTGLVGLGVEISRRLRSHNPRADPLVCATGLGS 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 408 AIFCLIFVAAKSIVGAYICIFVGTLLFSNWAITADILMYVIVIPTRATVALQSFTS 467
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 APFLSLACARSGSIVATYIFIGETLLSMNWAIVADILLYVVIPTRRSTAEAFQIVLS 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 468 HLLGDAGSPYLIGFISDLIRQKDSPLWEFLSIGVALMCPFVVVLGG 516
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 HLLGDAGSPYLIGLISDLRLRRNPPFLSEFRALQFSLMLCAFVGLGG 289

RESULT 2
US-09-149-476-555
; Sequence 555, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
```

US-09-149-476-554
Sequence 554, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 166 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 13.6%; Score 396; DB 4; Length 136;
Best Local Similarity 65.2%; Pred. No. 9.5e-28;
Matches 73; Conservative 19; Mismatches 20; Indels 0; Gaps 0;

QY 117 LDRYTVAGVLLDIOQHFQKVRGAGLQSVFVCSFMAPIRGYLGRNKRKVLSCGIF 176
DB 1 MDRFTVAGVLPDIOEFNGDSSGLIQTVFISSVWLPAPVFGYLDGRNKRKVLSCGIIA 60
QY 177 FMSVTFSSFPQQYFVLLVLSRGLVIGGEASYSTIAPTITGDLFTKNT 228
DB 61 FWSLVTLGSSFPGEHFVLLTLTRGLVGVGEASYSTIAPTITADLFPVADORT 112

RESULT 4

US-09-328-352-7272
Sequence 7272, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09-328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7272
LENGTH: 458
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7272

Query Match 8.6%; Score 250.5; DB 4; Length 458;
Best Local Similarity 24.2%; Pred. No. 5.3e-14;
Matches 111; Conservative 76; Mismatches 194; Indels 77; Gaps 17;

QY 110 LGNVLVNDRYTVAGVLLDIOQHFQKVRGAGLQSV-FICSPFWAAPIFGYLGRNKR 168
DB 37 LAVIFSDIRQIALMIEPIKADQLSDTQFSLHGLAPSLFYAVMGLPLAYIADFRSP 96
QY 169 VILSCGIFFWSAVT----FSSSFIPOQYFVLLVLSRGLVIGGEASYSTIAPTITGDLFTK 224
DB 97 KLISIGIIVWSLATATCGLSKNFIQ-----LFLSRMVGVGGAALSPAAYSFSDMFSK 150
QY 225 NRTMLSVFYPFALPGSLGYITGSSVKQ-----AAGDHWALRVSPVLG 270
DB 151 DKLGRAVGIYSIGAFLLGGGTAFLVGGYVNNLLKGVTLIEVPLGALKAKWQIAFLVGLPG 210
QY 271 MITGTILILV--PATKRGHADQLG--DQLKARTSWLRDMKALIRNRSYVFSLSATSAVS 326

DB 211 IILGLFILTVPKPARKGQQLNQSQVDQVKF--TQCLQFIKK--HAKTFACHYGLGFTFYA 267
QY 327 FATGALGMWIP-LYLHRAQVVQKTAETCNPPCAKSLFAGAITCTGTGFLGVVVTGAGAT 385
DB 268 MALYSLSWTPAFYIRKFLAP-----TETGYMLGTILLVANTLGVFCAGWLN 315
QY 386 RWCRLKTRADPLVCV-GMLGSAIFICLIFVAAKSSIVGAYICIFVGETLL-----PSN 439
DB 316 DMFIKKGRQDAPMFTGVIGV--LIPIAFTTQDQ-----LWLSVILLIIPAMPFAS 366
QY 440 W--AITADILMYVVIPTERRATAVALQSFTHLLGDAGSPYILGIFSLIROSTKDSPLWE 497
DB 367 FPLVISATALQMLAPNQFRARLSALFLVSNLIGLVGTTLVAITDKV-----415
QY 498 FLSLGYALMLCPFVVVLGGM--FFLATALFFVSDRARAE 534
DB 416 ---FQNLVWGSSLSIVGGLSCVLALALLFKGCKSFSE 450

RESULT 5

US-09-252-991A-30568
Sequence 30568, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30568
LENGTH: 352
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30568

Query Match 7.1%; Score 206.5; DB 4; Length 352;
Best Local Similarity 23.7%; Pred. No. 3.2e-10;
Matches 87; Conservative 56; Mismatches 141; Indels 83; Gaps 14;

QY 194 MLLVLSRGLVIGGEASYSTIAPTITGDLFTKNTRLMLSVFYFAIPLGSLGYITGSSV- 252
DB 1 WQFLTRFVGVGVGEAALSPAAYSLIADSPFRERRATAISVYSMGIYLSGLAFLGLGLVI 60
QY 253 --KQAGDWH-----WAL--RVSPVLGMITGTILILVLPATKRHADQLGDLQKAR 299
DB 61 KPASAQGDVHLPLFGEVFPWQLIFILGAAGVLFCLLLAIREPARRGVGAGVAVPLGEV 120
QY 300 TSWLRDMKALIRNRSYVFSLSATSATGALGMWIPLYLHRAQVVQKTAETCNSPPCG 359
DB 121 GAYLRANRKTVLCHNFGFACL-----SPAGYSGGAWPTFFVTRTHGWD-----A 164
QY 360 AKDSLIFGAICTCTGFLGVVTCAG-ATRCRLKTORADPLVCVAGMLGSAIFICLIFVAA 418
DB 165 GHGVVYGSIVAVFGCLGIVFGRLADYWA--KRGSDANM-RVGLL--AAWAVIFP---216
QY 419 KSSIVGAYICIFVGTLLF-----SNW--AITADILMYVVIP-----TRRA 457
DB 217 -----TLVYPLDNNANWAAALWAPTFFLSMPFGVAPAAIQEIMPNSMRG 261
QY 458 TAVALQSFTHLLGAGSPYILGIFSLIROSTKDSPLWEFLSLGLVALMCLCPVVVLGGM 517
DB 262 QASATYLFVVTILFGLGLGPTAVALVTFV-----FADDMALRYSLLLVTLAAVLGAV 313
QY 518 FLATLAL 524
DB 314 VLLIGL 320

RESULT 6

US-09-489-039A-13997
; Sequence 13997, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13997
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13997

Query Match 7.1%; Score 206.5; DB 4; Length 430;
Best Local Similarity 22.9%; Pred. No. 4.2e-10;
Matches 102; Conservative 82; Mismatches 200; Indels 61; Gaps 17;

Qy 94 PASLGRGRGAAATLSLGNVLYDRYTVAGVLLDIOQHFGVKGDRGAGLLQSFVICSFMV 153
Db 17 PLRIRRVQKITLMTFTAGVNFDRSLSVAGEAIRADIGLSATERGVLLSAFSLSYGF 76

Qy 154 AAPIFGYLGDRFNKRLGSCIGFWSAVTFSSSFIPOQYFWLLVLSRGLVIGESYSTI 213
Db 77 AQLPSGILLDRGLPRIVLGLAGLIFWSAQALTMGV-NSFSHFILLRIGL-GIGEAAPMPA 134

Qy 214 APTIIGDLFTYNTLMLSVFYFAIPGSLGLYITGSSVQAQAGDWALRVPVLGMIT 273
Db 135 GVKISINDWYAQREGTAVGFNSGTVLGOAIA-PPALVIMQLANGWMTFMVIGLAGIVV 193

Qy 274 GTLILILVPATKGHADQ-----LGDLQKAR-----TSWLRDMKALIRNRSYV 316
Db 194 GLCWYV-----GYRNRQFTQEEQQVLAEEAARPAKFSWL-----ALFKRTTW 242

Qy 317 FSSLATS AVSFATGALGMWIPYLHRAQVQVKTAECTNPPCGAKDSLIFGALITCFGTFL 376
Db 243 GMILGFGSVNYTGWLYTAWLPGLY-LQAQGLSLART-----GWAAIPFLAAAVGMVN 295

Qy 377 GVTGAGATR-WCKLTKQRADPLVCVAGMLGSAIFICLIFVAAKSSIVGAYI-----CIF 430
Db 296 GLVVDALARRGYDQAKTRKT---AIVIGLVLSALGTLVVQSSTPAQAVAFISNALFCVH 352

Qy 431 VGETLLFSNWAITADILMYVVIPTRRATAVALQSFTHLAGDAGSPVLIGFISDLIRQST 490
Db 353 FAGT---SAGM---LVQVWAEHKVASVAIQNFGSFVFA-SFAPIVTGWVVD-----TT 400

Qy 491 KDSPLWEFLSLG--YALMLCPFFVW 513
Db 401 HSNFALVIAAGTVTFAGALCYFFIV 425

RESULT 7

US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Gu, Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-130001
; CURRENT APPLICATION NUMBER: US/10/162.012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06

; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,363
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/227,068
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: PCT/US01/25475
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-162-012-46

Query Match 6.9%; Score 202.5; DB 4; Length 488;
Best Local Similarity 21.5%; Pred. No. 1.2e-09;
Matches 114; Conservative 88; Mismatches 195; Indels 133; Gaps 22;

Qy 95 ASLGRG-----RGAATAILSLGNVLYDRY---TVAGVLLDIOQHFGVKGDRGAGLLQS 145
Db 5 AALGGGFLFGYDGTGVIGGFLAL---IDFLFRGLLTSSGALAEI---VGYSTVLTGLWS 58

Qy 146 VFICSFVWAAPIFGYLGDRENK---VILSCGIFWMSAVFSS-----SFIPQOYFWLVLVS 199
Db 59 IFFLGRLLGSLFAGKLGDRFRKKSLIALVLEFVIGALLSGAAPGYTTIGLWAFYLLIVG 118

Qy 200 RGLVIGIGEASYSYTIATIGDLFTKTRTILMLSVFYFAIPGSLGLGYITGSSVKAAGD- 258
Db 119 RVLVGLGVGGASVJVPWYISEIAPKALRGALGSLYQLAITIGILVAAIIGLKNKTNDS 178

Qy 259 ----WFW-----ALRVSPVLGMITGTLILVLP-----ATKRG--HADQL 292
Db 179 ALMSWGWRIPLGLVLPALLLLGLLFLPESPRMLVEKGLKEAREVLAKLRGVEDVDQE 238

Qy 293 GDOLKA-----RTSWLRDMKALIRN---RSYVFSLSATSVAVSFATGALGMWIP 338
Db 239 IQEIKAELEATVSEKAGKAGWGELEFRGRTPKVRQRLMLGMVLMQAFQQLTG-----INA 293

Qy 339 YLHRAQVQVQKTAETCNPPCGAKDSLIFGALITCFGTGLGV----- 379

Db 294 IFYYSITFKSV-----GVSDSVASLLVTVGVNVFVTFVALIFLVDRFGRPLL 345
QY 380 -TGAGATRCRLKTORADPLVCAVGMGSAIFICLIFV-----AKSSIVGAYICIFVG 432
Db 346 LLGAGMAICFL-----ILGASIGVALLLNKPKDPSSKAAGIVAIVFILLF 392
QY 433 ETLTFSNWAIADILMYVVIPT-RATAVALOSFTSHLLGDAGSPVLIGFISDLIRQSTK 491
Db 393 IAFALGWGPIPWVILSELPFKVRSKALATAANWL-----ANFIIGFLFPYITG--- 444
QY 492 DSQLBEFLSLGVALMLCPFWVLGGM-FFLATALFFVSD-RARAEQHLGE 539
Db 445 -----AIGLALGGVFLVAGLLVLFILFVFFVFPETKGTLEIEE 486

RESULT 8

US-09-328-352-5768
; Sequence 5768, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5768
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5768

Query Match 6.8%; Score 197.5; DB 4; Length 408;
Best Local Similarity 21.6%; Pred. No. 2.5e-09;
Matches 99; Conservative 71; Mismatches 174; Indels 115; Gaps 17;

QY 105 AAILSGNLVLDRTYVAGVLD--IQHFGVKRGA--GLLOSVCFSFWVAIPFY 160
Db 16 AFVLMVGVVLAFFDKISIAVLFSDFHQSMSGIAEDKAKGLWMTSFLLAYGFSVFLSF 75
QY 161 LGDRFNKRVILSGCIEFFWAVTFSSSFIPOQYFWLVLVSRGLVGIGEASYSYTIPTIGD 220
Db 76 LGDIFNPKWLFWSVTSWGLMLCMGF-TTYSGLMLL-RVLGLAEGPLFALATYIVKQ 133
QY 221 LFTKNTRTLMLSVFYFAIFLGSGLGYITGSSVKQAAGDWHWALRVSPLVGMTITGLIL 280
Db 134 TVTDROQARASTWFLIGTGIGAFGLGPPITAAV-LAHHDWHTTFFVMAALTIALILSIVFG 192
QY 281 VPATKRGHADQLGDQKARTSW---LRDMKALIRNSYVFSSLSATSVSFATGALGMWIP 337
Db 193 LRNLQKTKTVELEGESK-RTNFKGHITANKVLVNSAFWLVCLEFNIALTYLWGLNSWVP 251
QY 338 LYLHRAQVQVQKTAETCNSPPCGAKDSLIFGATCTFTGLGVVTVGAGATRCRLKTORADP 397
Db 252 SYLMQDK-----GF-----NLK----- 263
QY 398 LVCVAGMLGSAIFICLIFVAKSSIVGAYICIFVGETLL--FSNWAITADILMYV-VIPT 454
Db 264 ---EFGVYSSFPFIAMLI---GEVVGAFLSDKLGRRATQVFGS-LLLAGIFMYVMVIMT 315
QY 455 RRATAVALOSF-----TSHLLGDAG-----SPVLIGF 481
Db 316 EPLLIIAAMSLSAMANGFVAAVALLARVTSNMGATAGGIFNGLGNFASAIAPVLIGY 375
QY 482 ISDLIRQSTKDSPLMBFLSLGVALMLCPFWVLGGMFFL 520
Db 376 I-----VMQTHSFNLGITLAAVAVIGSLFLV 402

RESULT 9

US-09-489-039A-9335

; Sequence 9335, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9335
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9335

Query Match 6.5%; Score 189.5; DB 4; Length 443;
Best Local Similarity 22.0%; Pred. No. 1.5e-08;
Matches 101; Conservative 72; Mismatches 200; Indels 87; Gaps 19;

QY 114 LNYLDRTYVAGVLDIQHFGVKRGAAGLQSVCFSFWVAIPFYGLDRFNKRVILSC 173
Db 29 VNYGDRATLSIAGTEVAKELGSLAVSMGYIFSAGWAYLLMQIPGGLLDKFGSKKVSY 88
QY 174 GIFPWSAVTFSSSFT---PQYFWL-LVLSRGLVGIGEASYSYTIPTIGDLFTKNTRTL 229
Db 89 SLFFWSLFTFLOGFDVDFPLAWAGVSMFFMFLGFSFSEAPFANARIVAAWFPKERT 148
QY 230 MLSVF---YFAIPLGSG-LGYITGSSVKQAAGDWHWALRVSPLVGMTITGLILVLPAT 284
Db 149 ASAIFNAQAQYFSLALFSPLLGLWT-----FALGWEHVFTVWGIIGFVL-TIIVKFEVN 201
QY 285 KRGHADQLGDQK-----ARTSWLRDMKALIRNSYVFSSLSATSAV 325
Db 202 PTDHPRMSAAELKYISEGAVVDMDHKKEATPAAGPKMDYIRQLLTNRMLGLVFFGQYPL 261
QY 326 SFATGALGMWIFLYL--HRAQVQKTAETCNSPPCGAKDSLIFGATCTFTGLGVVTVGAG 383
Db 262 NTITWFFLTWTFYLVQDKGWSILKVGVSAP-----ALFG---FAG---GVLGGLF 308
QY 384 AT---RWCRLKTORADPLVCAVGMGSAIFICLIFVAKSSIVGAYICIFVGETLLFSN 439
Db 309 SDYLIGRCCTLTFAKLPVLGM-LLASSIILC-NYTASTPLIVITLMALAPFGKFGALG 366
QY 440 WAITADILMYVVIPTRRATAVALOSFTSHLLGDAGS---PYLIGFISDLIRQSTKDSPLW 496
Db 367 WPVISDVAPKEI-----VGLCGGVNFVGNVASIATPLVIGYI-----VS 406
QY 497 EFLSLGYVALMLCPFWVLGG--MPFLATALFFVSDRAAE 534
Db 407 ELHSPNGAL-----IFVGGSAIIMMMVCLFVVGDIKRME 440

RESULT 10

US-09-252-991A-17560
; Sequence 17560, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17560
; LENGTH: 506

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17560

Query Match 6.4%; Score 188; DB 4; Length 506;
Best Local Similarity 20.1%; Pred. No. 2.4e-08;
Matches 121; Conservative 72; Mismatches 218; Indels 192; Gaps 23;
QY 25 RRR----RGAORGAGGCCGAGGAGGAAAGDEVTLSGVSRRAPTGPPTGPGC 80
DB 4 RRRAGFIRARRPAPASQ--SARRRVAARQARRGRSOTRSTPAARQNVMPDAPAS-- 58
QY 81 AATKAGCAQPKPASLGRGGAAGAAAILSGNVNLVLDRTYVAGVLDLQOHFGVKDORGA 140
DB 59 -----AABERLERPLSPYHRLVFVIAIAFFPDSMDLAMMTLLGSIKAEFLGSAQA 111
QY 141 GLLOSVCIFSMVAAPIFGVLDGRFNKRVLSGIFPWSAVTSSSIPQOYFWLLVLSR 200
DB 112 GLIASSFFGKVIAGALUSGMLADRFGRKPVQASIVLWGLASVLCSTAGD--LDSLTFYR 169
QY 201 GLVGIGEASYSTIAPTIIIGDLFTKNTRTLMLSVF-----YFAIPLGSGLG 245
DB 170 VLLGIGMGEPPIAQSLLSEMPASRRGKYIALMDGFPLGFVAAGCLSYFLPL----- 224
QY 246 YITGSSVKQAQDMHWALRVSPVLMGTGTLLILVLP-----ATKRGHADQLGDLK 297
DB 225 --TG-----WR-SIFLVLPALPAVFVLAIRFLIPESPRWLEQAGRRQADRV----- 267
QY 298 ARTSWLRDMKALI-----RNRSVFSSLSATSVPATGALGMWPL 338
DB 268 -----LRDIEARVNRSLGTELPLPQORERSRPGFFSAFAB-----LWSPA 311
QY 339 YLHRAQVVKTAETCNSPPCGAKDSLIFGATCTGFLGVVGTGAGATRCWLKTRADPL 398
DB 312 YRRRTQVW-----GLWFFAL---LGFYGI---TSWLSALLQSGGPA 347
QY 399 V-----CAVGMGLGSAIFC-----LIFVAAKSSIVGAYIC 428
DB 348 VTQSVYVTVLISLAGIPGLCAAWLVBESGKPKSCVLMLLGGGAMAVAYGTAVFGSLA 407
QY 429 IFVG-----ETLFSNVAITADILMYVIP-----TRATAVALQSTSHLGDAGSPYL 478
DB 408 LLIGFGLAMOFFLFGMWAV-----LYTTPELYPTSARATGSGFASAVGR-IGSLGLPLV 461
QY 479 IGFISDLIRQSTKDSPLWELSLGLYALMLCPFFVVVLGGMFFLATAL-----FFVSDRARAE 534
DB 462 TGLVLPITGGG-----VFTLGALCFGVAAALVWAFGIETRGRTL 501
QY 535 QHL 537
DB 502 EEL 504

RESULT 11
US-09-489-039A-10827
; Sequence 10827, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10827
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10827

Query Match 6.4%; Score 186; DB 4; Length 463;
Best Local Similarity 21.9%; Pred. No. 3.2e-08;
Matches 82; Conservative 68; Mismatches 165; Indels 60; Gaps 13;
QY 118 DRYTVAGVLDLQOHFGVKDORGAGLLOSVCIFSMVAAPIFGVLDGRFNKRVLSGIF 177
DB 50 DTAIGFIASDLVQEWGVEKSAALGPVNSAALVGLAVGALTAGPLADRIGRKKVILMSIV 109
QY 178 WSAVTSSSIPQOYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTKNTRTLMLSVF 237
DB 110 PGFESLLTATATS--LNQLTLLRFLTGLGGAAMPNATLMSYAPERRALLVNLMPVG 167
QY 238 IPIGSGIGYITGSSVKQAQDMHWALRVSPVLMGTGTLLILVLP-----ATKRGHADQ 291
DB 168 FPMGSSLGGLFSAWMIPIHYG-WQSVLVGLGVPMLLAVALIFLLPESVRYLVWVQHPAQ 226
QY 292 LGDQLKARTSWLRDMKALINRSYVSSLSATS--SPATGA-----L 332
DB 227 IAAILRRIAPLPAVEFVLRAGQVKEKSAIGVIFSPRYAVGTVMCLTYFMGLLI 286
QY 333 GMWIPLYLHR--AOWVQKTAETCNSPPCGAKDSLIFGATCTGFLGVVGTGAGATRCWL 390
DB 287 TSWLPLLRITGASMSQASITAIPLGCGGIVLILGL----- 325
QY 391 KTORADP-LYCAVGMGLGSAIFCII-FVAAKSSIVGAYICIFVGETLLFSNVAITA-DIL 447
DB 326 -MDKINPNKVAVGWLLTGVFVCLVGFSTSSALMG--VMVFIAGSIM--NGAQSSMPAL 380
QY 448 MYVVIPTR-RATAVA 461
DB 381 AAGFYPTQGRATGVA 395

RESULT 12
US-09-489-039A-10722
; Sequence 10722, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10722
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10722

Query Match 6.3%; Score 183.5; DB 4; Length 503;
Best Local Similarity 22.1%; Pred. No. 6e-08;
Matches 105; Conservative 62; Mismatches 204; Indels 105; Gaps 17;
QY 108 LSLGNVNLVLDRTYVAGVLDLQOHFGVKDORGAGLLOSVCIFSMVAAPIFGVLDGRPNR 167
DB 26 LLLWMLLSALDQDTIVALTPTIVGELGDLKLSWVV-TAYILSSTIAVPLYGRKFDLGR 84
QY 168 KVTL--SCGIFFWSAVTFSSSIPQOYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTKN 225
DB 85 KIVLVQAIGLFLVGSALCGLA-----QNMVQLVLMRGLGGLGGLVMSAAVADVIPPA 140
QY 226 TRTLMLSVFYFAIPLGSGIGYITGSSVKQAQDMHWALRVSPVLMGTGTLLILVLPATK 285
DB 141 NRGRYOGLFGGVFLGATVIGPLIGGLVQHA--SWRWIFVINLPLGLFALLIVGAVFHSSN 199
QY 286 RHADQLGDLKARTSWLRDMKALINRSYVSSLSATSVPAT---GALGMWIPLYLHR 342
DB 200 KRSQHQI-----DWL-----GAIVLSMALLCIILFTSEGGSVHAW----- 234

```

343 QY AQVVQKTAETCNPPCGAKODSLIFGAICTCTFGFLGVVTGAGATRWCKLKTQRADPL----- 398
235 DB -----NDP-----QLWCILAF-GIVGIIIGFYEERMAAEPIIPLALFR 271
399 QY -----VCAVGMGLGSAIFICILFVAAKSSI--VGAYICIFVGSTLLFSNWAIT 443
272 DB NRSFLCLSLIGFVIGMSLFGSVTFPLYLQVVKEATTEAGLQLIPLMGOLLTS--IIS 329
444 QY ADILMVVPIPTRATAVALQSGTSHLGDAGSPYLI GFISDLIRQSTKOSPMEFLUSLGY 503
330 DB -----SKTKYRLEFPIGLGVLGMV-----LLTRITIHSPLMQ----- 368
504 QY ALMLCPVVVLG-GMEFLATALFFVSDRARAEOHLGERRAGVRVYHQRGPGGTAL 558
369 DB ---LYLFTGVGLGAGLGVQVLVLAVQNAPMAQMYGATSGVILFRSIGSGIVAL 421

RESULT 13
US-09-492-709A-258
; Sequence 258, Application US/09492709A
; Patent No. 6720139
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001A
; CURRENT APPLICATION NUMBER: US/09/492,709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 258
; LENGTH: 444
; TYPE: PRT
; ORGANISM: E. Coli
US-09-492-709A-258

```

| | | | | |
|-----------------------|-------|--|-------|-------------------------------------|
| Query Match | 6.2% | Score 182; | DB 4; | Length 444; |
| Best Local Similarity | 21.2% | Pred. No. 6.9e-08; | | |
| Matches | 98; | Conservative | 69; | Mismatches 218; Indels 78; Gaps 16; |
| QY | 107 | ILSLGNVNLVDRYTVAGVLLDIQQHFGVKDGRGAGLLQSVFICFSFMAAPFVGLGDREN | 166 | |
| DB | 22 | IIFIVAVNYADRATLSIAGTEVAKELOLSAVMSGVIFSAFGWAYLLMQIPGGMLLDKFG | 81 | |
| QY | 167 | RKVILSCGIFPWSAVTFSSPIQ-QYFWL---LVLSRGLVIGIGEASYSTIAPTIIIGDLF | 222 | |
| DB | 82 | SKKVYSLFFWSLFTFLOQFVDVMPPLAWAGISHPFRFMLGFSEAPSPFANARIVAAMF | 141 | |
| QY | 223 | TKNTRTLMVSF---YFAIPLGSG-LGYITGSSVKQAAGDWHALRVSPVLGMITGTFLI | 277 | |
| DB | 142 | PTKERGTASAFNSAQYFSLALFSLGGLWLT-----FAWGWEHVFTVMGVIGFVLTTALW | 195 | |
| QY | 278 | LILVPATKRGHADQGDOLK-----ARTSWLRMKALINRSTVF | 317 | |
| DB | 196 | IKLI-HNPDPHMSAEELKPTISENGAVVMDHKKPGSAAASGPKLHYIKQLLSNRMLLG | 254 | |
| QY | 318 | SSLATSAVSPATGALCMWIPYLL--HRAQVQKTAETCNSPP-CGAKDSLIFGATICTFG | 374 | |
| DB | 255 | VFGQGYFINTTWFTLTFWPIYLVQEKGMSILKVLGVASIPALCGFAGVGLGV---FSD | 311 | |
| QY | 375 | FLGVVTVGAGATWCRLKTKQRADPLVCAGVMLGSAIFCLIFVAAKSSIVGAYICIFVGET | 434 | |
| DB | 312 | YL-----IKRGLSLTIAARKLPV--LGMLLASTIILCNYNTNTVLVMLMALAFFGKG | 362 | |
| QY | 435 | LLFSNWAITADILMVYIPTRRATVALQSFTHLLGDAGS---PYLIGFISDLIRQSTK | 491 | |

```

Db      363  FGALGWPISD-----TAPKEIVGLCGGVNFVGNVASIVPLVIGYL----- 400
Qy      492  DSPLEWFLSLGYALMLCFVVVLGGMFFLATATLFFVSDRARA 534
      :      :      :      :      :      :      :      :
Db      406  ---VSELHSFNAALVF----VGCSSALMAWVCYLFVVGDIKME 441
      :      :      :      :      :      :      :      :

RESULT 14
US-09-489-039A-12384
; Sequence 12384, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBS.
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12384
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12384

```

| | | | | | |
|-----------------------|--------------|--|---|-------------|----------|
| Query Match | 6.2%; | Score 180; | DB 4; | Length 442; | |
| Best Local Similarity | 22.2%; | Pred. No. 1e-07; | | | |
| Matches 104; | Conservative | 72; | Mismatches 190; | Indels 102; | Gaps 21; |
| Qy | 107 | ILSGNVLYLDR | TVVAGVLL-DIQOHFGVKDRGAGLLQSLVFCISFVMAAPIFGYLGDRF | 165 | |
| Db | 35 | LLSVGTWNYLDR | TLIGIVAPQLSKIEHIDPAMWGII | 93 | |
| Qy | 166 | NRKVILSGIFFWSAVTSSFIPOQYFWLLVLRGLVGVGEASYSTIAPTIIGDLFTKN | 225 | | |
| Db | 94 | GKMLTYALSIPFWSLFTLLOSF | TLGLKSLLLRLGL-GVSEAPCFPANSRIVSTWFPQH | 151 | |
| Qy | 226 | TRTLMLSVFYFPAI | PLGSGLVYITGSSVKQAAGDHWALRSPVLGMI | 273 | |
| Db | 152 | ERARATAT | -----YTVGEYIGLAA-----FSPLLFLELHHGWRTLPFLT | 191 | |
| Qy | 274 | GTLILILVPATK | -----GHADQLGDQLKAR-----TSWLDRMKALINRS | 314 | |
| Db | 192 | GGLGILFTLVWRRYFHPHESRTANQAELEYIGENSINNKIQNVFNN | ROARKLLGCRQ | 250 | |
| Qy | 315 | YVFESLA | -----TSAVGFATGALGMWITPLYLHRAQVVO--KTAETCTNSPPCGAKDSLI | 367 | |
| Db | 251 | ILGASLGQFAGNTLVLPFLT | -----WPFSYLANERHLPWLHVGFATWTPFLAAAILGILFG | 305 | |
| Qy | 368 | AITCFTGFLGVVTVGAGATWCRCLKTORADPLVCVAGMLGSAIFTCLFIPA | KSSIVGAYI | 427 | |
| Db | 306 | -----GWISDRLLKRTGSVNISRKLPIIS--GLLLSSCIITAAWNVSANSTVIIMS | 354 | | |
| Qy | 428 | CIFVGETLLFSNWAITADILVMYVVIPTERRATAV | ALOSFTSHLLGDAGSPYLIGFISDLI | 486 | |
| Db | 355 | VAFPGQGVGWGLWTIISDI | -----APENMAGLTGGIFNFCAN-MASIIAPLIIGVII | 405 | |
| Qy | 497 | RQSTKDSPLWBFSLGLEYALMLCPFFVVLGGMFFLATALFFV | VSDDRARAE | 534 | |
| Db | 406 | -----SATGNFF--YALIVIGLTALIGVAY | -----IFIIGDIKRIE | 440 | |

RESULT 15
US-09-711-164-428
; Sequence 428, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL


```
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711.164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 428
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-711-164-428

Query Match      6.0%; Score 176.5; DB 4; Length 425;
Best Local Similarity 22.7%; Pred. No. 2e-07;
Matches 98; Conservative 73; Mismatches 184; Indels 77; Gaps 20;

QY 110 LGNVLNLDRTYTAGVLLDIOQHFGVKDQKAGLLQSFVCSFVMAAPIFGVLGDRFRNKV 169
Db 42 LGYVDFGDFMFIFYLHIKADILGIDTIQATLIGTVAFIARPIGGGFFGAMADKYGRKP 101
QY 170 ILSGIFFWSAVTFSSSFIPQYFELLVLSRGLVIGEASVSTIAPTIIIGDLFTKNTRTL 229
Db 102 MWMWAIYISVGTGLSIATNLY--MLAVCRFIVGLGMSGEYACASTYAVESHPKNLQSK 159
QY 230 MLSVYFPAIPLGSGLVGTGSSVKQAAGDWHWALRVSPVLGMITGTLILILVPATKRGHA 289
Db 160 ASAPLVSGFSVGN--IIAAQIIPQFAEVYGM--RNSFFIGLLPVLVLWIRKSAPESQ- 213
QY 290 DQIGDQLKARTSWLRDMKALIRNSYVFSLSATSNVSF----ATGALGMWIPLYLHR--- 342
Db 214 EWIEDKYKDKSTFLS-----VFRKPHLSISMIVFLVCFLFGANWPINGLLPSYLDNGV 268
QY 343 AQVVQKTAETCNSPPCGAKDSLIFGAI--TCFTGFLGVVTGAGATRCWLKTKQADPLVC 400
Db 269 NTVVISLMTIAG-----LGLTGTIFFGFVGDKIGV-----KKAF 304
QY 401 AVGMLGSAIFIC-LIFVAAK--SIVGAYICIFVGETLLFSNWAITADI--LMYVVIPT- 455
Db 305 VVGLITSFIFLCPFFISVKNSSLIG--LCLF---GLMFTNLGIAGLVPKFIYDYPTKL 359
QY 456 RATAVALQSTSHLIGDAGSPYLIIGFISDLIROSTKDSPLWEFLSLGYALMLCPFVVVLG 515
Db 360 RGLGTGL-IYNLGATGMAAPVLATYISG-----YYGLGVSL----FIVTVA 401
QY 516 GMEFLATLFFV 527
Db 402 ---FSALLILLV 410
```

Search completed: October 23, 2004, 13:16:46

Job time : 28 secs

This Page Blank (uspto)

Sat Oct 23 17:45:11 2004

```

; ORGANISM: Homo sapiens
US-10-085-198-48

Query Match      100.0%; Score 2920; DB 15; Length 566;
Best Local Similarity 100.0%; Pred. No. 8.4e-216;
Matches 566; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNCLECAAAAGAAEEEADEARRRRRGAORAGAGCGCCGARGAGAGVSAAGDEVQTL 60
   |||
Db 1 MNCLECAAAAGAAEEEADEARRRRRGAORAGAGCGCCGARGAGAGVSAAGDEVQTL 60

QY 61 SSVRRAPTPGPTGCTPCCATAGKAGQAPKPSLGRGCAAAAILSLGNVLNLYDRY 120
   |||
Db 61 SSVRRAPTPGPTGCTPCCATAGKAGQAPKPSLGRGCAAAAILSLGNVLNLYDRY 120

QY 121 TVAGVLLDIIQQHFGVKDRGAGLLQSVFICSFVMAAPIFGYLGDRFNKRVILSCGIFFWSA 180
   |||
Db 121 TVAGVLLDIIQQHFGVKDRGAGLLQSVFICSFVMAAPIFGYLGDRFNKRVILSCGIFFWSA 180

QY 181 VTFSSSFIPOQYFWLLVLSRGLVIGEASYSIIAPTIIIGDLFTKTRTILMLSVFYFAIPL 240
   |||
Db 181 VTFSSSFIPOQYFWLLVLSRGLVIGEASYSIIAPTIIIGDLFTKTRTILMLSVFYFAIPL 240

QY 241 GSGLGVIITGSSVKQAAGDMHVALRVSPVLGMITGTLILVLPATKRGHADQLGDQLKART 300
   |||
Db 241 GSGLGVIITGSSVKQAAGDMHVALRVSPVLGMITGTLILVLPATKRGHADQLGDQLKART 300

QY 301 SWLRDMKALIRNRSYVFFSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNSPPCGA 360
   |||
Db 301 SWLRDMKALIRNRSYVFFSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNSPPCGA 360

QY 361 KDSLIIFGATCTFTGFLGVVTGAGATRCWLKTQADPLVCVCMGLSAIFICLIFFVAKS 420
   |||
Db 361 KDSLIIFGATCTFTGFLGVVTGAGATRCWLKTQADPLVCVCMGLSAIFICLIFFVAKS 420

QY 421 SIYGAVICIFVGTLLFSNWAITADILMVYVITPRATAVALQSFTSHLLGDAGSPYLIG 480
   |||
Db 421 SIYGAVICIFVGTLLFSNWAITADILMVYVITPRATAVALQSFTSHLLGDAGSPYLIG 480

QY 481 FISDLIRQSTKDSPLMEFLSLGYALMLCPVVVLGGMFFLATALFFVSDRARAEOHLGER 540
   |||
Db 481 FISDLIRQSTKDSPLMEFLSLGYALMLCPVVVLGGMFFLATALFFVSDRARAEOHLGER 540

QY 541 RAGVRVHHQRPGRPGGTALAHRVVGAS 566
   |||
Db 541 RAGVRVHHQRPGRPGGTALAHRVVGAS 566
```

```

RESULT 2
US-10-168-651-9
; Sequence 9, Application US/10168651
; Publication No. US2003017125A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dyoung Aina M.
; APPLICANT: YANG, Junming
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameen R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT APPLICATION NUMBER: US/10/168,651
; CURRENT FILING DATE: 2002-06-21
```

```

; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758;
; 60/181,625
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
; 2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 3232992CD1
US-10-168-651-9

Query Match      67.7%; Score 1978; DB 14; Length 398;
Best Local Similarity 99.0%; Pred. No. 1.3e-143;
Matches 384; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 152 MVAAPIFGCLGDRFNKRVILSCGIFFWSAVTFSSSFIPOQYFWLLVLSRGLVIGEASYS 211
   |||
Db 1 MVAAPIFGCLGDRFNKRVILSCGIFFWSAVTFSSSFIPOQYFWLLVLSRGLVIGEASYS 60

QY 212 TIAPTIIGDLFTKTRTILMLSVFYFAIPLGSGLGVIITGSSVKQAAGDMHVALRVSPVLGM 271
   |||
Db 61 TIAPTIIGDLFTKTRTILMLSVFYFAIPLGSGLGVIITGSSVKQAAGDMHVALRVSPVLGM 120

QY 272 ITGTLLILVLPATKRGHADQLGDQLKARTSWLRDMKALIRNRSYVFFSSLSATSAVSFATGA 331
   |||
Db 121 ITGTLLILVLPATKRGHADQLGDQLKARTSWLRDMKALIRNRSYVFFSSLSATSAVSFATGA 180

QY 332 LGMWIPLYLHRAQVQVQKTAETCNSPPCGAKDSLIFGATCTFTGFLGVVTGAGATRCWLK 391
   |||
Db 181 LGMWIPLYLHRAQVQVQKTAETCNSPPCGAKDSLIFGATCTFTGFLGVVTGAGATRCWLK 240

QY 392 TORADPLVCVCMGLSAIFICLIFFVAKSSIVGAYICIFVGTLLFSNWAITADILMVYV 451
   |||
Db 241 TORADPLVCVCMGLSAIFICLIFFVAKSSIVGAYICIFVGTLLFSNWAITADILMVYV 300

QY 452 IPRTRATAVALQSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLMEFLSLGYALMLCPPV 511
   |||
Db 301 IPRTRATAVALQSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLMEFLSLGYALMLCPPV 360

QY 512 VVLGGMFFLATALFFVSDRARAEOHLGE 539
   |||
Db 361 VVLGGMFFLATALFFVSDRARAEOHVNQ 388
```

```

RESULT 3
US-10-024-623-5
; Sequence 5, Application US/10024623
; Publication No. US20020187524A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Roy A. J.
; TITLE OF INVENTION: 8099, 46455, 54414, 53763, 67076, 67102, 44181.
; TITLE OF INVENTION: 67084FL, AND 67084 ALT, HUMAN PROTEINS AND METHODS OF
; FILE REFERENCE: MNI-214CP
; CURRENT APPLICATION NUMBER: US/10/024,623
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-024-623-5
```

Query Match 44.5%, Score 1298; DB 13; Length 528;
Best Local Similarity 53.4%; Pred. No. 3.1e-91;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

Qy 53 AGDEVQTLSSVRRAPTGP-PGTGTPGCAATAGKGAQPKPASLGR-----GRGA-A 104
Db 2 AGSDTAPLQADDDPGVPVGTGPGTGNPKSBEPEVDQGLQRTGLSPGRSALI 61

Qy 105 AAILSLGNVLNLDYRTVAGVLLDIOQHFGYKDRGAGLQSVFICSPFWAAPIFGYLGDR 164
Db 62 VAVLCYNLLNMDRFTVAGVLPDIEQFFNIGDSSGLIQTVFISSYVWLAPVFGYLGDR 121

Qy 165 FNRKVIILSCGIFFWASVTFSSFPQQYFWLLVLSRGLVGEASYSYTIAPTIIGDLFTK 224
Db 122 YNRKYLCCGGIAFWLSVTLGSSFPGEHFWLLLLTRGLVGVGEASYSYTIAPTLIADLFA 181

Qy 225 NTRTLMLSVFFAIPGLSGLYINGSSVKQAAGDHWALRSPVLGMITGTLLIILVPAT 284
Db 182 DQSRMLSIIFYFAIPVSGLYTAGSKVKQWAGDHWALRVTGPGVAVLLFLVVRP 241

Qy 285 KRGHADQLGDLKAR-TSWLRDMKALIRNRSYFSSLSATSAVSFATGALGMWIPLYLHRA 343
Db 242 PRGAVERHSDLPPLNPTFSWADLRALARNPSFVLSLGTAVAFVTSGLAWAPAFLLRS 301

Qy 344 QVQKTAETC-NSPPCGAKOSLIFGATICTGTFGLGVVGTAGATRWCLKTQRADPLVCV 402
Db 302 RVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHNSPRADPLVCAT 361

Qy 403 GMLGSAIFCICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462
Db 362 GLGSAAPFLSLACARGSIYATYIFIFIGTLLSMNWAIVADILLYVVIPTRRSTAEF 421

Qy 463 QSFTSHLLGDAGSPYLIGFTSDLIROSTKDSPLMEFLSLGVALMCLCPVVVLGGMFFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMCLCAFVGLGGAFLGT 481

Qy 523 ALFPVSDRAEAOHLGE--RRAGVR-----VVHQRG 551
Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 516

RESULT 4
US-10-154-419-55
; Sequence 55, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria A.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 55
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-55

Query Match 44.5%, Score 1298; DB 14; Length 528;
Best Local Similarity 53.4%; Pred. No. 3.1e-91;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

Qy 53 AGDEVQTLSSVRRAPTGP-PGTGTPGCAATAGKGAQPKPASLGR-----GRGA-A 104
Db 2 AGSDTAPLQADDDPGVPVGTGPGTGNPKSBEPEVDQGLQRTGLSPGRSALI 61

Qy 105 AAILSLGNVLNLDYRTVAGVLLDIOQHFGYKDRGAGLQSVFICSPFWAAPIFGYLGDR 164
Db 62 VAVLCYNLLNMDRFTVAGVLPDIEQFFNIGDSSGLIQTVFISSYVWLAPVFGYLGDR 121

Qy 165 FNRKVIILSCGIFFWASVTFSSFPQQYFWLLVLSRGLVGEASYSYTIAPTIIGDLFTK 224
Db 122 YNRKYLCCGGIAFWLSVTLGSSFPGEHFWLLLLTRGLVGVGEASYSYTIAPTLIADLFA 181

Qy 225 NTRTLMLSVFFAIPGLSGLYINGSSVKQAAGDHWALRSPVLGMITGTLLIILVPAT 284
Db 182 DQSRMLSIIFYFAIPVSGLYTAGSKVKQWAGDHWALRVTGPGVAVLLFLVVRP 241

Qy 285 KRGHADQLGDLKAR-TSWLRDMKALIRNRSYFSSLSATSAVSFATGALGMWIPLYLHRA 343
Db 242 PRGAVERHSDLPPLNPTFSWADLRALARNPSFVLSLGTAVAFVTSGLAWAPAFLLRS 301

Qy 344 QVQKTAETC-NSPPCGAKOSLIFGATICTGTFGLGVVGTAGATRWCLKTQRADPLVCV 402
Db 302 RVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHNSPRADPLVCAT 361

Qy 403 GMLGSAIFCICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTTRATAVAL 462
Db 362 GLGSAAPFLSLACARGSIYATYIFIFIGTLLSMNWAIVADILLYVVIPTRRSTAEF 421

Qy 463 QSFTSHLLGDAGSPYLIGFTSDLIROSTKDSPLMEFLSLGVALMCLCPVVVLGGMFFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMCLCAFVGLGGAFLGT 481

Qy 523 ALFPVSDRAEAOHLGE--RRAGVR-----VVHQRG 551
Db 482 AIFTEADRRRAQLHVQGLLHEAGSTDDRIIVVPQRG 516

RESULT 5
US-10-146-733-50
; Sequence 50, Application US/10146733
; Publication No. US20030165891A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria A.
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; FILE REFERENCE: MNI-248
; CURRENT APPLICATION NUMBER: US/10/146,733
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/185,938
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/515,520
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/518,866
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/195,993
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/199,799
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 60/233,537
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/235,018
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/235,059
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 09/796,720

FILE REFERENCE: D0273 NP
CURRENT APPLICATION NUMBER: US/10/648,593
CURRENT FILING DATE: 2003-08-26
PRIOR APPLICATION NUMBER: 60/406,385
PRIOR FILING DATE: 2002-08-27
NUMBER OF SEQ ID NOS: 557
SOFTWARE: PatentIn version 3.2
SEQ ID NO 251
LENGTH: 528
TYPE: PRT
ORGANISM: Homo sapiens
US-10-648-593-251

Query Match 44.5%; Score 1298; DB 16; Length 528;
Best Local Similarity 53.4%; Pred. No. 3.1e-91;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLGSGVRRAPTGP-PGTPGTCGAATAGKGAQPKPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPGVPGTGNPKSEPEVDPQEGLOQITGLSPGRSALI 61
QY 105 AAILSLGNVLYDRYTVAGVLLDIQOHFGVKDRGAGLLQSVFICSFMAAIPFGYLGDR 164
DB 62 VAVLCYINLLNMDRFTVAGVLPDIEQFFNIGDSSSGLIQTFFISSYMLAPVFGYLGDR 121
QY 165 FNRKVLSCGIFPWSAVTFSSFIPOQYFWLLVLSRGLVGVGEASYSYTIPTIIGDLFTK 224
DB 122 YNRKYLKMGCGIAFWSLVTLGSSFIPEGHFHLLLLTRGLVGVGEASYSYTIPTIADLFVA 181
QY 225 NTRTLMLSVFFPAIPGLSGLYGTGSSVKQAAGDHWALRVSPVLGMITGTLLILVLPAT 284
DB 182 DQSRMLSFYFAIPVGSGLGYTAGSKVKDMAGDHWALRVTPGLGVVAVLLFLVVRP 241
QY 285 KRGHADQLGDLKAR-TSWLRDMKALIRNSYVFSLSLATSVAVSFATGALGMWIPLYLHRA 343
DB 242 PRGAVERHSDLPPLNPTSMWADRLARNPSEVLSLGFATAVAVTGSALWAPAFLLRS 301
QY 344 QVVKQTAETC-NSPPCCGAKDSLIFGAICTFTGFLGVVGTGAGATRCWLKTKQADPLVCV 402
DB 302 RVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHSNPRADPLVCAT 361
QY 403 GMLGSAIFICLIFFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVAL 462
DB 362 GLLGSAPFLFSLACARGSIYATYIFIFIGETLLSNWAIADILLYVVIPTRRSTAEAF 421
QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFFVVLGGMFFLAT 522
DB 422 QIVLSHLLGDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMLCAFVGALGGAFLGT 481
QY 523 ALFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
DB 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIWVVPQRG 516

RESULT 7
US-10-403-571-52
Sequence 52, Application US/10403571
Publication No. US20040068763A1
GENERAL INFORMATION:
APPLICANT: Hopkins, Nancy
APPLICANT: Golling, Gregory
APPLICANT: Amsterdam, Adam
APPLICANT: Sun, Zhaoxia
TITLE OF INVENTION: Developmental Mutations in Zebrafish
FILE REFERENCE: 01997/539002
CURRENT APPLICATION NUMBER: US/10/403,571
CURRENT FILING DATE: 2003-03-25
PRIOR APPLICATION NUMBER: US 60/368,760
PRIOR FILING DATE: 2002-03-29
NUMBER OF SEQ ID NOS: 159
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 52
LENGTH: 507

PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 09/828,035
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 09/833,081
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: US 09/843,128
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: US 09/957,683
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: US 09/964,252
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 09/964,256
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US 10/024,623
PRIOR FILING DATE: 2001-12-17
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 50
LENGTH: 528
TYPE: PRT
ORGANISM: Homo sapiens
US-10-146-733-50

Query Match 44.5%; Score 1298; DB 14; Length 528;
Best Local Similarity 53.4%; Pred. No. 3.1e-91;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLGSGVRRAPTGP-PGTPGTCGAATAGKGAQPKPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPGVPGTGNPKSEPEVDPQEGLOQITGLSPGRSALI 61
QY 105 AAILSLGNVLYDRYTVAGVLLDIQOHFGVKDRGAGLLQSVFICSFMAAIPFGYLGDR 164
DB 62 VAVLCYINLLNMDRFTVAGVLPDIEQFFNIGDSSSGLIQTFFISSYMLAPVFGYLGDR 121
QY 165 FNRKVLSCGIFPWSAVTFSSFIPOQYFWLLVLSRGLVGVGEASYSYTIPTIIGDLFTK 224
DB 122 YNRKYLKMGCGIAFWSLVTLGSSFIPEGHFHLLLLTRGLVGVGEASYSYTIPTIADLFVA 181
QY 225 NTRTLMLSVFFPAIPGLSGLYGTGSSVKQAAGDHWALRVSPVLGMITGTLLILVLPAT 284
DB 182 DQSRMLSFYFAIPVGSGLGYTAGSKVKDMAGDHWALRVTPGLGVVAVLLFLVVRP 241
QY 285 KRGHADQLGDLKAR-TSWLRDMKALIRNSYVFSLSLATSVAVSFATGALGMWIPLYLHRA 343
DB 242 PRGAVERHSDLPPLNPTSMWADRLARNPSEVLSLGFATAVAVTGSALWAPAFLLRS 301
QY 344 QVVKQTAETC-NSPPCCGAKDSLIFGAICTFTGFLGVVGTGAGATRCWLKTKQADPLVCV 402
DB 302 RVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLRHSNPRADPLVCAT 361
QY 403 GMLGSAIFICLIFFVAAKSSIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVAL 462
DB 362 GLLGSAPFLFSLACARGSIYATYIFIFIGETLLSNWAIADILLYVVIPTRRSTAEAF 421
QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFFVVLGGMFFLAT 522
DB 422 QIVLSHLLGDAGSPYLIGLISDRLRNWPSPFLSEFRALQFSLMLCAFVGALGGAFLGT 481
QY 523 ALFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
DB 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIWVVPQRG 516

RESULT 6
US-10-648-593-251
Sequence 251, Application US/10648593
Publication No. US20040106132A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS

[illegible]

```
Db      2 AGSDTAPFLSQADDDPDPGVPCTGPGTGNPKSBEPEVPDQEGQLQRTGLSPGRSALI 61
Qy      105 AAILSGNVLYLDRTYTVAGVLLDIOOHFGVKDRGAGLLQSVFICSEFWAAPIFGVLGDR 164
Db      62 VAVLCYINLLNMDRFTVAGVLDLEQFFNIGDSSGLIQTVFISSYMWLAPVFGVLGDR 121
Qy      165 FNRKVLSCGIFPWSAVTFSSSFIPOQYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTK 224
Db      122 YNRKYLCCGIAFWSLVTLGSSFIPEGEHFWLLLLTRGLVGVGEASYSTIAPTIIADLFVA 181
Qy      225 NTRTLMLSVYFAIPALPGSLGYITGSSVKQAAGDHWALRVSPVLGMITGTILILVPA 284
Db      182 DQSRMLSIIFYFAIPVGSGLGYIAGSKVKMAGDHWALRVTPGLGVAVALLFLVVRP 241
Qy      285 KRGHADQLGDLKAR-TSWLRDMKALIRNRSYVFSLSATSVSFATGALGMWIPLYLHRA 343
Db      242 PRGAVERHSDLPPLNPTSWADLRARN----- 270
Qy      344 QVVQKTAETCNSPPCGAKDSLIFGAITCFTGFLGVVTVGAGATRWCRCLKTORADPLVC 403
Db      271 -----LIFGLITCLTGLVGLGVEISRRLHNSNPRADPLVCATG 310
Qy      404 MLGSAIFICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMVVVIPTTRATAVALQ 463
Db      311 LLGSAPFLFLSLACARGSVATYIFIFIGETLLSMNWAIVADILLYVVIPTRRSTAEAFQ 370
Qy      464 SFTSHLLGDAGSPYILIGFISDLIRQSTKDSPLWEFLSLGYALMCLPCFVVVLGGMFLATA 523
Db      371 IVLSHLLGDAGSPYILIGLISDLRLRNWPPSPFSEFRALQFSLMCLCAFGALGGAFLGTA 430
Qy      524 LFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
Db      431 IFIEADRRRAQLHVQGLLHEAGSTDDRIVVVPORG 464
```

```
RESULT 10
US-10-169-395-126
; Sequence 126, Application US/10169395
; Publication No. US20040034192A1
; GENERAL INFORMATION:
; APPLICANT: KATO, Seishi
; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAS ENCODING
; FILE REFERENCE: 01997.015100.US
; CURRENT APPLICATION NUMBER: US/10/169,395
; PRIOR FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: JP 2000-585
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-588
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-2299
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-26862
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: JP 2000-58367
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/JP00/09359
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 150
; SEQ ID NO 126
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-169-395-126

Query Match      40.0%; Score 1167; DB 15; Length 476;
Best Local Similarity 49.4%; Pred. No. 3.2e-81;
Matches 254; Conservative 53; Mismatches 141; Indels 66; Gaps 7;

Qy      53 AGDEVOTLSGVSRRAPTGP-PGTPPGCAATAGKGAQPKPASLGR-----GRGA-A 104
Db      2 AGSDTAPFLSQADDDPDPGVPCTGPGTGNPKSBEPEVPDQEGQLQRTGLSPGRSALI 61
```

RESULT 11

```
US-09-860-232A-2
; Sequence 2, Application US/09860232A
; Patent No. US20020028494A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; TITLE OF INVENTION: 57256 AND 58289, NOVEL HUMAN
; FILE REFERENCE: 381552001500
; CURRENT APPLICATION NUMBER: US/09/860,232A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,288
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-860-232A-2

Query Match      37.1%; Score 1084; DB 9; Length 512;
Best Local Similarity 47.5%; Pred. No. 8.3e-75;
Matches 223; Conservative 75; Mismatches 163; Indels 8; Gaps 3;

Qy      75 PGTPGCAATAGKGAQPKPA-----SLGRGRG-AAAAISLGNVLYLDRTYTVAGVLLD 128
Db      12 FPGGGLQGQSPGPGRCQCPPTPTTSWSLPPNRAVYVAAVLCYINLLNMYNMFIIAGVLLD 71
Qy      129 IQOHFGVKDRGAGLLQSVFICSPFWAAPIFGVLDGRFNRKVLSCGIFWWSAVTSSSFI 188
Db      72 IQEVPQISDNHAGLLQTVFVSCLLSAFVFGVLDGRHSRKATMSFGIILLWSGAGSSSFI 131
Qy      189 PQOYFWLLVLSRGLVGIGEASYSTIAPTIIIGDLFTKTRTLMLSVYFAIPALPGSLGYIT 248
Db      132 SPRVSWLFLSRGIVGTGSASYSTIAPTIVGLDFVRDQTRVLAVFYIFIPVGSGLGYVL 191
Qy      249 GSSVQKAAAGDHWALRVSPVLGMITGTILILVPAIKRGHADQLGDLKA--RTSWLRDM 306
```

Db 192 GSAVTMLTGNWRWALRVMPCLAEVALILLILLVDPDPRGAAETQGEAGVGFRRSWCEDV 251
Qy 307 KALIRNRSYVFFSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNPSPPCGAKDSLIF 366
Db 252 RYLGKWSFVWSTLGVMTAMAVTTCALGFMAKPFLEARVHGLQPPCFQPCSNPDSLIF 311
Qy 367 GAITCTFTGLVGVVGTGATRCWLKTRADPLVCAVGMLSAIFICLIFVAAKSSIVGAY 426
Db 312 GALTMTGTVIGVILGAEASRRYKVPICGAEPLICASSLLATAPCLYLALVLAPTTLLASY 371
Qy 427 ICIFVGTGTLFSSNWAITADILMYVVIPTTRATAVALQSFTHLIGDAGSPYLIGFISDLI 486
Db 372 VFLGLGELLSSCNWAVVADILLVVVPRCGTAEALQITVGHILGDAGSPYLIGLSSVL 431
Qy 487 RQSTKDSPLWELSLGVALMCLPFFVVLGGMFFLATALFFVSDRARAQ 535
Db 432 RARRPDSYLQFRSLQOSFLCCAFVIALGGCFLLTALYLERDETRAWQ 480

RESULT 12

US-10-297-022-23
; Sequence 23, Application US/10297022
; Publication No. US20030216310A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: THORNTON, Michael
; APPLICANT: WALIA, Narinder K.
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: LAL, Preeti
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOLEY, Catherine M.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: TANG, Y. Tom
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YANG, Junming
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: SANJANWALA, Madhu S.
; APPLICANT: RAUMANN, Brigitte E.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: GREENE, Barrie D.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: KEARNEY, Liam
; APPLICANT: ELLIOTT, Vicky S.
; APPLICANT: SEILHAMER, Jeffrey J.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: BURFORD, Nell
; APPLICANT: DING, Li
; APPLICANT: LU, Dyrng Aina M.
; APPLICANT: HILLMAN, Jennifer L.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: FI-0109 PCT
; CURRENT APPLICATION NUMBER: US/10/297,022
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: 60/208,424; 60/209,001; 60/210,588; 60/212,335; 60/213,747;
; PRIOR FILING DATE: 2000-05-26; 2000-06-01; 2000-06-08; 2000-06-16; 2000-06-22; 2000-06-26
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 23
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030216310A1 1713377CD1
US-10-297-022-23

Query Match 37.1%; Score 1084; DB 14; Length 512;
Best Local Similarity 47.5%; Pred. No. 8.3e-75;
Matches 223; Conservative 75; Mismatches 163; Indels 8; Gaps 3;
Qy 75 PGTGCAATAGKAGCAQPKPA-----SLGRGRG-AAAAILSLGNVLYNLYDRYTVAGVLLD 128
Db 12 PGPGLGQGGPGFGRCQPPITPTSLPPWRAVAAAVLCYNLLNMMFIIAGVLLD 71
Qy 129 IQOHFGVKDRGAGLQSVFICSMVAAPIEFYGLGDRNRKRVILSCGFFVFSVAVTFSSFI 188
Db 72 IQEVFQISDNHAGLLQTVFVSCILLSSAPVFGYGLDRHSRKATMSFGILLHSGAGLSSFI 131
Qy 189 PQYFVLLVLSRGLVGIGEASYSYTIPTIIGDLFTKNTRTLMLSVFYFAIPLGSGLYIT 248
Db 132 SPRYSWFLFLSRGIVGTGSASYSTIAPTVLGDLFVRDQRTVLAVFYFIPVGSGLGYVL 191
Qy 249 GSSVKQAGDWHWALRVSPVLGMITGTLIIILLIPATKRGHADQLGDOLKA--RTSWLRDM 306
Db 192 GSAVTMLTGNWRWALRVMPCLAEVALILLILLVDPDPRGAAETQGEAGVGFRRSWCEDV 251
Qy 307 KALIRNRSYVFFSSLSATSAVSFATGALGMWIPLYLHRAQVQVQKTAETCNPSPPCGAKDSLIF 366
Db 252 RYLGKWSFVWSTLGVMTAMAVTTCALGFMAKPFLEARVHGLQPPCFQPCSNPDSLIF 311
Qy 367 GAITCTFTGLVGVVGTGATRCWLKTRADPLVCAVGMLSAIFICLIFVAAKSSIVGAY 426
Db 312 GALTMTGTVIGVILGAEASRRYKVPICGAEPLICASSLLATAPCLYLALVLAPTTLLASY 371
Qy 427 ICIFVGTGTLFSSNWAITADILMYVVIPTTRATAVALQSFTHLIGDAGSPYLIGFISDLI 486
Db 372 VFLGLGELLSSCNWAVVADILLVVVPRCGTAEALQITVGHILGDAGSPYLIGLSSVL 431
Qy 487 RQSTKDSPLWELSLGVALMCLPFFVVLGGMFFLATALFFVSDRARAQ 535
Db 432 RARRPDSYLQFRSLQOSFLCCAFVIALGGCFLLTALYLERDETRAWQ 480

RESULT 13

US-09-826-734-160
; Sequence 160, Application US/09826734
; Publication No. US20030017457A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Mishra, Vishnu S.
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Kekuda, Ramesha
; TITLE OF INVENTION: Novel Polynucleotides and Polypeptides Encoded Thereby
; FILE REFERENCE: 15966-754
; CURRENT APPLICATION NUMBER: US/09/826,734
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,576
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 270
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160
; LENGTH: 193
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-826-734-160
Query Match 33.4%; Score 976; DB 10; Length 193;
Best Local Similarity 98.4%; Pred. No. 5.2e-67;
Matches 190; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 143 LQSVFICSMVAAPIEFYGLGDRNRKRVILSCGFFVFSVAVTFSSFPQQYFVLLVLSRGL 202
Db 1 LPAVFCSEFMVAAPIEFYGLGDRNRKRVILSCGFFVFSVAVTFSSFPQQYFVLLVLSRGL 60
Qy 203 VGIGEASYSYTIPTIIGDLFTKNTRTLMLSVFYFAIPLGSGLYITGSSVKQAAGDWHWA 262

Db 61 VIGGEASYSTIAPTIIIGDLFTKNTRLMLSVFYFAIPGLSSGLGYITGSSVKQAAGDHWHA 120
QY 263 LRVSPVLGMITGTLILILVPATKRGHADQDGLKARTSWLRDMKALINRYSYVFFSSLAT 322
Db 121 LRVSPVLGMITGTLILILVPATKRGHADQDGLKARPSWLRDMKALINRYSYVFFSSLAT 180
QY 323 SAVSFATGALGMW 335
Db 181 SAVSFATGALGMW 193
RESULT 14
US-10-154-419-81
; Sequence 81, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckman, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255salt MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 81
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-154-419-81
Query Match 29.3%; Score 855; DB 14; Length 531;
Best Local Similarity 40.0%; Pred. No. 3.4e-57;
Matches 194; Conservative 82; Mismatches 173; Indels 36; Gaps 9;
QY 71 PPGTPTP-----GCAATAGPQAQPPKASLGRGRGAAAILSGNVLYLDRTY 121
Db 23 PPYTTPTDSPEDKIRSNSTATTASQPEFQ-----GCWTIVVAILFIINLLNYMDRTY 75
QY 122 VAGVLDDIQHFGVKDQAGLQSVFICSMVAAPIFGYLGRFNRKVLSCGIFPWSAV 181
Db 76 IAGVLNDVQTYNISDAWAGLIQTTPWFVFIIFSPICGFLGRYNNRKFVVGIAIWSA 135
QY 182 TFSSSIPQYFWLLVLSRGLVGIGEASSTIAPTIIIGDLFTKNTRLMLSVFYFAIPGLG 241
Db 136 VFASIPISNQFWLFLFRGIVGIGEASVAIISPTVIADMTGVLRSRLMWFYFAIPFG 195
QY 242 SGLGYITGSSVQAAGDHWALRVSPVLGMITGTLILILVPATKRGHADQDGLKA--- 298
Db 196 CGLGFVVGSAVSWTGHQWQVRVTGVLGIVCLLIIVFVRSPERGKAEREKGEIAAASPE 255
QY 299 RTSWLRDMKALINRYSYVFFSSLATSAVSATGALGMWIPLYLHRAQVVKTAETCNSPPC 358
Db 256 ATSYLDMDKDLISNATYVSSISGTYATVFMVGTAWMADITIQYADSAERNGTITEDOK- 314
QY 359 GAKDSLIFGAITCFTGFLGVVWGA-CATRWCR-----LKTQADPLVCVAGMLGSAIF 410
Db 315 -ANINLVFALNLCVGVGLVAGTLVSNMWSRGVGPFGKHQIQTIVRADALVCAI---GAALC 370
QY 411 ICLIFVA---AKSSIVGAYICIFVGETLLFSNWAITADILMVVPIPTRTATVALQSF 467
Db 371 IPTLILAIQNIENNNFANGMLFICIVASSFNWATNVLDLLSVVPPQRSSASSWQLIS 430
QY 468 HLLGDAGSPYLGTGFTSDLRQSTKSPLEFSLGVALMCPVVVVLGCMFFLATALFV 527
Db 431 HMFPGASGYIILGLISDAIR-GNEDTAQAHYKSLVTSFWLCVGTLLVLSILFGISAITVV 489
QY 528 SDRAR 532

Db 490 KDKAR 494
RESULT 15
US-10-146-733-31
; Sequence 31, Application US/10146733
; Publication No. US20030165891A1
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Gluckmann, Maria A.
; APPLICANT: Silos-Santiago, Inmaculada
; TITLE OF INVENTION: NOVEL TWIK-6, TWIK-7, IC23927, TWIK-8, IC47611, IC47615,
; TITLE OF INVENTION: HNMDA-1, TWIK-9, alpha2delta-4, 54414, AND 53763 MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: MNI-248
; CURRENT APPLICATION NUMBER: US/10/146,733
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/185,938
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/515,520
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/518,866
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: US 60/195,734
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 60/195,993
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 60/199,799
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: US 60/233,537
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: US 60/235,018
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/235,059
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 60/256,240
; PRIOR FILING DATE: 2000-12-15
; PRIOR APPLICATION NUMBER: US 60/256,588
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/258,028
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 09/796,720
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/828,035
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 09/833,081
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 09/843,128
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 09/957,683
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 09/964,252
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 09/964,256
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 10/024,623
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 31
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-733-31
Query Match 29.3%; Score 855; DB 14; Length 531;
Best Local Similarity 40.0%; Pred. No. 3.4e-57;
Matches 194; Conservative 82; Mismatches 173; Indels 36; Gaps 9;
QY 71 PPGTPTP-----GCAATAGPQAQPPKASLGRGRGAAAILSGNVLYLDRTY 121
Db 23 PPYTTPTDSPEDKIRSNSTATTASQPEFQ-----GCWTIVVAILFIINLLNYMDRTY 75

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 12:50:48 ; Search time 27 Seconds
(without alignments)
2016.988 Million cell updates/sec

Title: US-10-085-198-48

Perfect score: 2920
Sequence: 1 MMCLCASAAGGAEBEAD.....VHQRGPGGTALAHRRVVGAS 566

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79.*

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|----------|---------------------|
| 1 | 855 | 29.3 | 531 | 2 T19232 | hypothetical prote |
| 2 | 815 | 27.9 | 488 | 2 T19854 | hypothetical prote |
| 3 | 681 | 23.3 | 411 | 2 T20623 | hypothetical prote |
| 4 | 392.5 | 13.4 | 507 | 2 B84616 | hypothetical prote |
| 5 | 354.5 | 12.1 | 746 | 2 T05899 | hypothetical prote |
| 6 | 285 | 9.8 | 438 | 2 G87290 | major facilitator |
| 7 | 278 | 9.5 | 457 | 2 R31064 | probable MFS trans |
| 8 | 276 | 9.5 | 519 | 2 E87557 | major facilitator |
| 9 | 252 | 8.6 | 453 | 1 P62250 | hypothetical 49.4 |
| 10 | 252 | 8.6 | 453 | 2 D91293 | probable transport |
| 11 | 252 | 8.6 | 453 | 2 G86134 | hypothetical prote |
| 12 | 247 | 8.5 | 444 | 2 D87557 | major facilitator |
| 13 | 219 | 7.5 | 546 | 2 S48313 | hexose transport p |
| 14 | 217 | 7.4 | 479 | 2 F87474 | major facilitator |
| 15 | 214 | 7.3 | 422 | 1 A69853 | hexuronate transpo |
| 16 | 205 | 7.0 | 541 | 1 MMBYH2 | glucose transport |
| 17 | 204.5 | 7.0 | 436 | 2 AG0272 | probable integral |
| 18 | 203.5 | 7.0 | 457 | 2 D71144 | hypothetical prote |
| 19 | 203 | 7.0 | 422 | 2 D72302 | hypothetical prote |
| 20 | 200 | 6.8 | 456 | 1 S27616 | probable glucarate |
| 21 | 196 | 6.7 | 472 | 2 B99372 | multidrug resistanc |
| 22 | 195.5 | 6.7 | 422 | 2 H69839 | multidrug resistanc |
| 23 | 195 | 6.7 | 490 | 2 E96810 | probable transport |
| 24 | 194 | 6.6 | 455 | 1 H69752 | probable glucarat |
| 25 | 194 | 6.6 | 458 | 1 YTESY8 | tetracycline resis |
| 26 | 192 | 6.6 | 477 | 2 B75409 | multidrug-efflux t |
| 27 | 192 | 6.6 | 472 | 2 F90355 | multidrug resistanc |
| 28 | 183.5 | 6.3 | 418 | 2 S76669 | hypothetical prote |
| 29 | 183.5 | 6.3 | 459 | 2 S42238 | tetracyclin resist |

| | | | | | |
|----|-------|-----|-----|----------|-----------------------|
| 30 | 183.5 | 6.3 | 521 | 2 T50023 | hypothetical prote |
| 31 | 182.5 | 6.2 | 537 | 2 D45634 | hypothetical prote |
| 32 | 182 | 6.2 | 372 | 2 D75094 | transport protein, re |
| 33 | 182 | 6.2 | 412 | 2 S39734 | chloramphenicol re |
| 34 | 182 | 6.2 | 444 | 2 E85974 | probable galactara |
| 35 | 182 | 6.2 | 444 | 2 E91129 | probable galactara |
| 36 | 182 | 6.2 | 444 | 2 C65102 | probable galactara |
| 37 | 182 | 6.2 | 450 | 2 A65061 | probable glucarate |
| 38 | 181.5 | 6.2 | 586 | 2 B87408 | tetracycline resis |
| 39 | 181 | 6.2 | 433 | 2 G86812 | D-xylose proton-ey |
| 40 | 181 | 6.2 | 452 | 2 AE0861 | probable glucarate |
| 41 | 181 | 6.2 | 458 | 1 YTESRT | tetracycline resis |
| 42 | 180.5 | 6.2 | 458 | 2 S23743 | tetracycline resis |
| 43 | 180.5 | 6.2 | 567 | 2 S31294 | hexose transport p |
| 44 | 179.5 | 6.1 | 372 | 2 H71070 | hypothetical prote |
| 45 | 178 | 6.1 | 421 | 2 AD3393 | bicyclomycin resis |

ALIGNMENTS

RESULT 1

T19232

hypothetical protein C13C4.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T19232

R:Harris, B.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19095

A:Accession: T19232

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-531 <WIL>

A:Cross-references: UNIPROT:O01927; EMBL:Z92825; PIDN:CAB07311.1; GSPDB:GN00023; CESP:

A:Experimental source: clone C13C4

C:Genetics:

A:Gene: CESP:C13C4.5

A:Map position: 5

A:Introns: 78/1; 145/3; 269/2; 341/3; 389/3; 411/3; 500/1

Query Match 29.3%; Score 855; DB 2; Length 531;
Best Local Similarity 40.0%; Pred. No. 1.5e-50;
Matches 194; Conservative 82; Mismatches 173; Indels 36; Gaps 9;

| | | | |
|----|-----|--|-----|
| Qy | 71 | PGTPTGTP-----GCAATAGCGAQQPKPASLGRGCGAAAILSLGNVLNLYDRYT | 121 |
| Db | 23 | PPYTPPTDSPEDKIRSNSTATTASQPEFQ-----GCWTIVVVAIIIFIINLLNMYDRYT | 75 |
| Qy | 122 | VAGVLLDIQQHFQVKRGAGILQSVFICSFMAAIPFGYLGDRFNKRVILSCGIFFWSAV | 181 |
| Db | 76 | IAGVLNDVQYNNISDAWAGLIQTFMVFFIIFSPICGFLGDRYRKWIFVVGIAIWISA | 135 |
| Qy | 182 | TFSSSFIPQYFWLLVLSRGLVGIGEASYSIAFTIIGDLFTKNTRLMLSVFFPAIPLG | 241 |
| Db | 136 | VFASTPIPSNQWFLFLFRGIVGIGEASYSIAISPTVIADMTFTGLVRSRLMVFVFAIPFG | 195 |
| Qy | 242 | SLGIYITGSSVKQAGDWHALRVSPVLGMITGTLLILLVPATKRGHADQGDQLKA--- | 298 |
| Db | 196 | CGLGFVVGSASVASTGHWQMGVRVTGLVIGVLCILLIIVFVREPERGKABREKGAASAE | 255 |
| Qy | 299 | RTSLRLDMKALIRNRSYVFFSSLSATSVSFATGALGMWIPLYLHRAQVQVQVQVQVQVQV | 358 |
| Db | 256 | ATSVLDNKKLLSNATVTSLSGYTATVFWVGTLAWPATIQVADSARRNGTITEDOK- | 314 |
| Qy | 359 | GAKDSLIFGATCTFGTGLVWVGA-GATRCR-----LKTORADPLVCVAVGLGSAIF | 410 |
| Db | 315 | -ANINLVFGALTCVGVGLVGAIGTFLVSNMWSRGVGPFKHIQTVRADALVCAI---GA | 370 |
| Qy | 411 | ICLIFVA---AKSIVGAYICIFVGETLLSPNWAITADILMYVIVPTTATVALQSF | 467 |
| Db | 371 | IPTLLAIQNIENSNMFWAGMLFCIVASSFNWATNVDLLSVVVPQRRSSASSWQILIS | 430 |

Db 10 PPVTETTSRCYSTSSSTPLABLETVRSLEIVVESSSLSPPVLLVIFCIINLLNMDRGAI 69
Qy 123 AG-----VLLDQQHFGVDRGAGLIQSIFVIFCSFVMAAFIFGYLGRF 165
Db 70 ASNGVNGSTRSCNDKCKTATGQGHENLSPEDGVLSSEFMVGLLIASIFASLAKR- 128
Qy 166 NRKVLSCGIFPWSAVTF--SSSFIPOQYFWLLVLSRGLVGIGBASVSTIAPTIGLFT 223
Db 129 ----LIGVGLTWTIIVLGCSSFA---FWFVLCRMFVGVGSEAFISLAAPFIDNAP 180
Qy 224 KNTETMLSVFPALPGSGLYITGSSVKQ-----AAGDWHALRVSP--VLGMITGLI 277
Db 181 QEQKAAMGLFMYCIPSGVALGYVGVGHFNRVAFGEAVIMAPFAVLGIMPLQ 240
Qy 278 L-----ILVPATKRGHADQLGQDKARTSWLRDMKALINR 313
Db 241 LKGSSETLKNNLQVDNEIHDQFEVSIETSKSYANAV---FKSFTGFAKMKVLYKEK 297
Qy 314 SYFPSSLATSAVSFATGALGMWIP-----LYLHRAQVQKTAETCNSPPCGAKDSLIFGA 368
Db 298 VFVNVVLGYVSNFVIGAYSYWGPKAGYNIY-----KMKQAD-----MIFGA 339
Qy 369 ITCTGFLGVVGTGAGATWCLKTQADPLVCAVGMGSAIFCLIFVAAKSSIVGAYIC 428
Db 340 VTIICGVITLSSGGFILDRTVATIPNAPKLLSGATFLG-AVF-CFTAFTLKS--LYGFIA 395
Qy 429 IF-VGETLLFSNMAITADILMYVVIPTRRATAVALQSFTHLGDAGSPYLIGFISLIR 487
Db 396 LFALGELLVATQAPVNVVCHCVKPSLRPLSMAISVVAHIFGDEVSSPLVGLVQDHIN 455
Qy 488 QSTKDSPLWFLSLGYALMLCPFVVVLGGMFFLATALFFVS-----DRARAEQ 535
Db 456 SWRKT-----LILTSILFLAAIWFIGIFINSVDRFNOE 491

RESULT 5
T05899
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05899
R:Bevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schon, O.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, April 1998
A:Reference number: 215456
A:Accession: T05899
A:Molecule type: DNA
A:Residues: 1-746 <BEV>
A:Cross-references: UNIPROT:O49546; EMBL:AL021684
A:Experimental source: cultivar Columbia; BAC clone F6H11
C:Genetics:
A:Map position: 5
A:Introns: 54/2; 99/2; 134/2; 198/3; 235/3; 265/2; 309/2; 360/2; 395/2; 420/3; 445/3; 48
A:Note: F6H11.180

Query Match 12.1%; Score 354.5; DB 2; Length 746;
Best Local Similarity 24.5%; Pred. No. 1.8e-16;
Matches 119; Conservative 72; Mismatches 185; Indels 109; Gaps 15;

Qy 107 ILSGNVNLVDRTVA-----GVL---LDIQHFGVDRGAGLIQSIFIC 149
Db 269 ILCIINLINVDRGVIASNGVSSKVCDAKGVCSAGTIGQGFNLTFEDGLSSAFV 328
Qy 150 SFVMAAPIFVGLGDRNRK-----VILSCGIF-----FWSAVTFSSSFIPOQY 192
Db 329 GLLVASPIFAGLSKRFNYQQHFVFLFFGVFNPFKLGVLTVWTIIVGCGF--SYN 386
Qy 193 FWLLVLSRGLVGIGSEASYTIAPTIIIGDFTKNTRLMLSVFYFAIPLGSLGYITGSSV 252
Db 387 FWMIAVFMFVGGEASFISLAAPYIDDSAPVARKNFWLGLFYNCIPAGVALGVVFGYI 446
Qy 253 KQAGDWHALRVSPVIGMTGTILILVLPATK-RGHADQ-----LGD 294
Db 447 GNHLG-WRWAFYIEAIAMAVFVLSFCIKPPQQLKGFADKSKKPSSTIETVAPTDAEAS 505

Qy 295 QLKARTS-----WLRDMKALINRNSYVSSLSATSAVSFATGALGMWIP-----LYLH 341
Db 506 QIKTKTPKSNLVVLFGKDLKALPSEKVFIVNVGLYITVNFVIGAYSYWGPKAGFYI-- 563
Qy 342 RAQVQKTAETCNSPPCGAKDSLIFGAICTFTGFLGVVGTGAGATWCLKTQADPLVCA 401
Db 564 ----KMKQAD-----MIFGGLTICGIIIGLGGSYVLDNRNATLSNTFKLAA 607
Qy 402 VGMGSAIFCLIFVAAKSSIVGAYICIF-VGETLLFSNMAITADILMYVVIPTRRATAV 460
Db 608 STLGA-----FCTAFMKMNAFIALFAVEGELIFAPQAPVNFVCLHCVRNLRPLSM 663
Qy 461 ALOSFTSHLGDAGSPYLIGFISDLIRQSTKDSPLWFLSLGYALMLCPFVVVLGGMFFL 520
Db 664 ASSTVLHILGDPSSPLYGKMDHLKWRKST-----LIITSILFL 705
Qy 521 ATALP 525
Db 706 AAIW 710

RESULT 6
G87290
major facilitator family transporter CC0336 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87290
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87290
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-438 <STO>
A:Cross-references: UNIPROT:Q9AB95; GB:AE005673; NID:gl3421483; PIDN:AAK22323.1; GSPDB
C:Genetics:
A:Gene: CC0336

Query Match 9.8%; Score 285; DB 2; Length 438;
Best Local Similarity 25.3%; Pred. No. 5.2e-12;
Matches 112; Conservative 73; Mismatches 220; Indels 38; Gaps 14;

Qy 93 KPASLGRG-----RGAAAAILSGNVNLVDRTVAGVLL-DIQHFGVDRGAGLIQSIF 147
Db 12 KAETAGRGGRYRYVVLAMLILAYTFNFLDR-QILGILAGSIKAEHLTDTQLGLMGVAF 70
Qy 148 ICSFMVAAPIFGVLGDRFNKVLISCGIFFWSAVTFSSSFIPOQYFVLLVLSRGLVGIGE 207
Db 71 AALYTLGVPLAMLADRVSRKTWIMTVALTVWSGFTVVCGLAGG--FWSLFLARMGVGE 128
Qy 208 ASYSTIAPTIIIGDFTKNTRLMLSVFYFAIPLGSLGYITGSSVKAQAGDWHALRVSP 267
Db 129 AGGVAPAYSLLIADYFPEQRARALAVVSGIPLGTALGVLFGLI-AAVVDWRFATVAG 187
Qy 268 VLGMITGTLILVLPATKRGHADQL-GDQLKAR-----TSWLRDMKALINRNSYVSSLSAT 322
Db 188 LAGVAFAPIFKVVVKOPVRGLDRAPGEVAPAPPPKAPAFQVLAATVMPKPSFWLLSFGA 247
Qy 323 SAVSFATGALGMWIPLYLHRAQVQKTAETCNSPPCGAKDSLIFGAICTFTGFLGVVGTGA 382
Db 248 ACSICGYVAFVLPITFFQKSFGLSLDR-----ALYTSALSLFGGVAGIWWGG 296
Qy 383 GATRWCLKTQADPLVCAVGMGSAIFCLIFVAAKSSIVGAYICIFV--GETLLFSN 440
Db 297 VLADRGAKNAAYALAPALCFL--VALPCFLLAMNVQSMVAFLFLIPTGLNLAMLG 354
Qy 441 AITADILMYVVIPTRRATAVALOFTSHLGDAGSPYLIGFISDLIRQSTKDSPLWFLS 500
Db 355 VWAA--VOHLAPPSMRTTTSALFLLINLLGLAVGLWFFGVVSDLL-----TPRYGAES 406

QY 501 LCVYALML-CPFVVVLGGMFFLAT 522
Db 407 MRYAIYYGUSFYVAAVZLILAS 429

RESULT 7

F83064
probable MFS transporter PA4654 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83064
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950, MUID:20437337, PMID:10984043
A:Accession: F83064
A:Molecule type: DNA
A>Status: preliminary
A:Cross-references: UNIPROT:Q9HVD8; GB:AE004879; GB:AE004091; NID:g9950901; PIDN:AAG0804
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4654

Query Match 9.5%; Score 278; DB 2; Length 457;
Best Local Similarity 23.7%; Pred. No. 1.6e-11;
Matches 108; Conservative 77; Mismatches 185; Indels 86; Gaps 16;
QY 106 AILSLGNVLYDRYTVAGVLLDIQHFVGVKDRGAGLQSVFICSPWVAAP1-FGVLDGR 164
Db 19 AILMAVYLSFIDRQITLNLVGVIRDLAISDTMSLMGLSFALPYTCGPIGLGRMAN 78
QY 165 FNRKVLSCGIFWSAVTFSSPIPOQYFVLVLSRGLVGEASYSYTIPTIIGDLFTK 224
Db 79 RSRGILFLGVLSWNTAACGL--ARSYMQFLTFRVGVGVEALSPAYSIIADSFR 136
QY 225 NTRTLMVYFPAIFLGSLGYITGSSV---KQAGDMH-----WAL--RVSPVLG 270
Db 137 ERRATAISVYSGIYGLGLFLGLVLFKFAAQGDVHLPLFGEVRPQLIFLILGAAG 196
QY 271 MITGTILILVPAKRGHADQLKARTSWLRDMKALIRNSYVFSLSATSAVSFATG 330
Db 197 VLFCLLLLAIREPARGVGAGVAVPLGEVGYLURANRKTVLCHNFGFACL-----SFAGY 251
QY 331 ALGMWIPLYLHRAQVQKTAETCNSPPCGAKDSLIFGAICTCTGFLGVVTGAG-ATRMCR 389
Db 252 GSGAWPTFFVRTHGD-----AGHVGVVYGSIVAVFGCLGIVFGGLADYWA- 299
QY 390 LKTQRADPLVCVGMGLGSAIFICLIPVAAKSSIVGAYICIFVGETLLF-----SNW--AI 442
Db 300 -KRGSDANN-RVGLL--AAWAVIP-----TLVYPLLDNANWAAAL 337
QY 443 TADILMVVIP-----TTRATAVALQSTSHLLGDAGSPVLI GFISDLIRQ 488
Db 338 MAPTVFELSPFGVAPAAIQEIMPNSMRGQAGAIYLVVTLFGLGVGPTAVAVLTDV-- 395
QY 489 STKDSLWELSLGLYALMLCPFFVVVLGGMFFLAT 524
Db 396 -----FADMALRYSLLVLTAAVNLGAVVLIGL 425

RESULT 8
E87557
major facilitator family transporter CC2486 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: E87557
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, J.C.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249, MUID:21173698, PMID:11259647
A:Accession: E87557
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-519 <STO>
A:Cross-references: UNIPROT:Q9A5G4; GB:AE005673; NID:g13424039; PIDN:AAK24457.1; GSPDB:1
C:Genetics:
A:Gene: CC2486

Query Match 9.5%; Score 276; DB 2; Length 519;
Best Local Similarity 24.3%; Pred. No. 2.5e-11;
Matches 111; Conservative 71; Mismatches 188; Indels 86; Gaps 14;
QY 82 ATAKPGQAQPKPASPGRGAAAAILSL---GNVLNVLDRYTVAGVLLDIQHFVGVKDR 138
Db 24 AENHASGGDRPLYSN-----GYKATVLLGLLATYTFNFIDRTIITIGQAIKVDLKLTD 78
QY 139 GAGLLQSVFICSPWVAAP1FGV---LGDENRKRKVLSCGIFWSAVTFSSSFIPOQVFW 194
Db 79 QLGLLGGY---FALLYTLIGIPARLAERFNRVTIISVSLVMSG--FTALCGAAANPA 133
QY 195 LLVLSRGLVGEASYSYTIPTIIGDLFTKNTRLMLSVFVPAIFLGSLGYITGSSVKQ 254
Db 134 QLALFRFGVGEAGCSPSHSLISDYEPKRAPALSIYSGIPLGTWFGAVAGWLQA 193
QY 255 AAGDWHWALRVSPVLGMITGTLLILVLPATKRGHA-----DQLGDQLKARTSWLR 304
Db 194 EFSRWFVAVVGLPGLLAVIVKLWVKEPPRGHSEIVERPLEAEDVVVPEAKPAFMSMAN 252
QY 305 DMKALIRNSYVFS-----LATSAVSFATGALGMWIPLYLHRAQVQKTAETCNSPP 357
Db 253 EFKELMWATKILFGKPVLMHVGLVGTIASFGAYSGAFVPSYFVRAFDLGL----- 303
QY 358 CGAKDSLIFGAICTCTGFLGVVTGAGATRCRLKTQRADPLVCVGMGLGSAIFICL-IFV 416
Db 304 --AAGVLTGLIGFSAGVGLVGGFLSDWAGKRSKAKVALTPAIGLI-----LCTPIVI 356
QY 417 AAKSSIVGAYICIFVGETLLFSNWAITADILM-----YV-----VIPRRAT 458
Db 357 LA-----YLOQDMQTALLILVPGIFHYVVLAPTFFGVVQNSVEPRRRAT 400
QY 459 AVALQSTSHLLGDAGSPVLI GFISDLIRQSTKQSP 494
Db 401 ATALLFFFLNIALGVGPTFGVLTGWLIDHLAQLHFNNP 436

RESULT 9

F85250
hypothetical 49.4 kD protein in ter-mdoB intergenic region - Escherichia coli (strain K-
C:Species: Escherichia coli
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: F85250; S56582
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: F85250
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-453 <BLAT>
A:Cross-references: UNIPROT:P9398; GB:AE000506; GB:U00096; NID:g2367377; PIDN:AAC77312
A:Experimental source: strain K-12, substrain MG1655
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92
A:Reference number: S56344; MUID:95334362; PMID:7610040
A:Accession: S56582
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 262-453 <BUR>

A;Cross-references: EMBL:U14003; NID:g12631172; PIDN:AAA97253.1; PID:g537198
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C;Superfamily: hexuronate transporter

```
Query Match      8.6%; Score 252; DB 1; Length 453;
Best Local Similarity 24.3%; Pred. No. 9.3e-10;
Matches 115; Conservative 79; Mismatches 212; Indels 68; Gaps 19;

Qy 72 PGTGTPGCAATAGP--GAQPKPASLGRGRGAAAAIISLGNVNLNLDRTYVAGVLLDI 129
Db 10 PRSFTPSSADIPVPDGLVQ-RSTRIKRIQTAMLLFFAAVINYLDRSSLSVANLTI 68

Qy 130 QOHFGVKDRGAGLLQSVFICFMAVAPIFGYLGDRFNKRVILSCGIFFWSAVTSSFP 189
Db 69 REELGLSATEIGALLSVFSLAYGIAQLPCGPLLDRKGRPRMLGLGMFFWSLFOAMSGMV- 127

Qy 190 QOYFWLLVLSRGLVIGEASYSTIAPTIIGDLFTKNTRTLMLSVFYFAIPLGSLGYITG 249
Db 128 -HSFTQFVLVRIGMIGEAPMPCGVKVINDFWNIKERGRPMGFNAASTIGVAVSPPI 186

Qy 250 SSVKQAAGDWHWALRVSPVLG--MITGLTIL-----ILVPATKRGHADQLDQKARTS 301
Db 187 AAMMLVNG-WRGMPITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 302 WL--RDMKALIRNSYVFSLSATSAVSFATGALGMWIPLYLHRAQVQVKTAEATCNSPPG 359
Db 244 PLGFAEWRSLFRNRTMGMMGLGFSGINYTAWLYLAWLPGLYQTAYNLD----- 291

Qy 360 AKDSLIFGATCTGTFGLGVVTVGAGATRW-----CRILKTORADPLVCVAV-GMLGSAIFIC 412
Db 187 AAMMLVNG-WRGMPITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 413 LIFVAAKSS-----IVG-AYICIFVGETLLFSNWAITADILMYVIVPRTATAV-ALQSET 466
Db 348 IVPQATTSMTAVLLIGNALFCIHAGT---SCWG-----LIHVAVASRMTASVGSIQNFA 399

Qy 467 SHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFVVVLGGMFFL 520
Db 400 SFICA-SFAPITIGFIVDTTH-----SFRLLALICGCVTAAGALAYI 440
```

RESULT 10

D91293
probable transport protein Ecs5316 [imported] - Escherichia coli (strain O157:H7, substra

C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: D91293
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene

A;Reference number: A9629; MUID:21156231; PMID:11258796

A;Accession: D91293

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-453 <HAY>

A;Cross-references: UNIPROT:O8XEG8; GB:BA000007; PIDN:BA38739.1; PID:g13364794; GSPDB:G

A;Experimental source: strain O157:H7, substrain RMD 0509952

C;Genetics:

C;Superfamily: hexuronate transporter

```
Query Match      8.6%; Score 252; DB 2; Length 453;
Best Local Similarity 24.3%; Pred. No. 9.3e-10;
Matches 115; Conservative 79; Mismatches 212; Indels 68; Gaps 19;
```

Qy 72 PGTGTPGCAATAGP--GAQPKPASLGRGRGAAAAIISLGNVNLNLDRTYVAGVLLDI 129

Db 10 PRSFTPSSADIPVPDGLVQ-RSTRIKRIQTAMLLFFAAVINYLDRSSLSVANLTI 68

Qy 130 QOHFGVKDRGAGLLQSVFICFMAVAPIFGYLGDRFNKRVILSCGIFFWSAVTSSFP 189

Db 69 REELGLSATEIGALLSVFSLAYGIAQLPCGPLLDRKGRPRMLGLGMFFWSLFOAMSGMV- 127

```
Qy 190 QOYFWLLVLSRGLVIGEASYSTIAPTIIGDLFTKNTRTLMLSVFYFAIPLGSLGYITG 249
Db 128 -HSFTQFVLVRIGMIGEAPMPCGVKVINDFWNIKERGRPMGFNAASTIGVAVSPPI 186

Qy 250 SSVKQAAGDWHWALRVSPVLG--MITGLTIL-----ILVPATKRGHADQLDQKARTS 301
Db 187 AAMMLVNG-WRGMPITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 302 WL--RDMKALIRNSYVFSLSATSAVSFATGALGMWIPLYLHRAQVQVKTAEATCNSPPG 359
Db 244 PLGFAEWRSLFRNRTMGMMGLGFSGINYTAWLYLAWLPGLYQTAYNLD----- 291

Qy 360 AKDSLIFGATCTGTFGLGVVTVGAGATRW-----CRILKTORADPLVCVAV-GMLGSAIFIC 412
Db 187 AAMMLVNG-WRGMPITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 413 LIFVAAKSS-----IVG-AYICIFVGETLLFSNWAITADILMYVIVPRTATAV-ALQSET 466
Db 348 IVPQATTSMTAVLLIGNALFCIHAGT---SCWG-----LIHVAVASRMTASVGSIQNFA 399

Qy 467 SHLLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCPFVVVLGGMFFL 520
Db 400 SFICA-SFAPITIGFIVDTTH-----SFRLLALICGCVTAAGALAYI 440
```

RESULT 11

G86134

hypothetical protein yjiz [imported] - Escherichia coli (strain O157:H7, substrain EDL9

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C;Accession: G86134

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimmalanta, E.; Potamousis, K.; Apodaca

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: G86134

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-453 <STO>

A;Cross-references: UNIPROT:O8XEG8; GB:AB005174; NID:g12519379; PIDN:AAG59539.1; GSPDB:

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

C;Gene: yjiz

C;Superfamily: hexuronate transporter

```
Query Match      8.6%; Score 252; DB 2; Length 453;
Best Local Similarity 24.3%; Pred. No. 9.3e-10;
Matches 115; Conservative 79; Mismatches 212; Indels 68; Gaps 19;
```

Qy 72 PGTGTPGCAATAGP--GAQPKPASLGRGRGAAAAIISLGNVNLNLDRTYVAGVLLDI 129

Db 10 PRSFTPSSADIPVPDGLVQ-RSTRIKRIQTAMLLFFAAVINYLDRSSLSVANLTI 68

Qy 130 QOHFGVKDRGAGLLQSVFICFMAVAPIFGYLGDRFNKRVILSCGIFFWSAVTSSFP 189

Db 69 REELGLSATEIGALLSVFSLAYGIAQLPCGPLLDRKGRPRMLGLGMFFWSLFOAMSGMV- 127

Qy 190 QOYFWLLVLSRGLVIGEASYSTIAPTIIGDLFTKNTRTLMLSVFYFAIPLGSLGYITG 249

Db 128 -HSFTQFVLVRIGMIGEAPMPCGVKVINDFWNIKERGRPMGFNAASTIGVAVSPPI 186

Qy 250 SSVKQAAGDWHWALRVSPVLG--MITGLTIL-----ILVPATKRGHADQLDQKARTS 301

Db 187 AAMMLVNG-WRGMPITIGVLGIFLAIGWYMLYRNREHVELTAVEQAYLN--AGSVNARD 243

Qy 302 WL--RDMKALIRNSYVFSLSATSAVSFATGALGMWIPLYLHRAQVQVKTAEATCNSPPG 359

Db 244 PLGFAEWRSLFRNRTMGMMGLGFSGINYTAWLYLAWLPGLYQTAYNLD----- 291

Qy 360 AKDSLIFGATCTGTFGLGVVTVGAGATRW-----CRILKTORADPLVCVAV-GMLGSAIFIC 412

Db 292 LKSTGLMAAIPFLFGAAGMLVNGVYTDVLVKGGMPIKSRK---ICIIAGMFCSSAAFTL 347
QY 413 LIFVAAKSS-----IVG-AVICFVGETLLFSNWAITADILMYVITPRATAV-AIQST 466
Db 348 IVPQATTSMMVALLIGMALCFIHFACT--SCWG-----LIHVAVASRMTASVGSIQNPA 399
QY 467 SHLLGDAGSPYLIGTSLIRQSTKDSPLWEFLSLGVALMCPVFFVVLGGMPEFL 520
Db 400 SFICA-SFAPITGFIYDTH-----SFLALIICGCVTAAGALAYI 440

RESULT 12
DB7557
major facilitator family transporter CC2485 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: DB7557
R:Niekman, W.C.; Feldbiyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n.; J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: DB7557
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: UNIPROT:Q9A5G5; GB:AE005673; NID:G13424038; PIDN:AAK24456.1; GSPDB:G
C:Genetics:
A:Gene: CC2485

Query Match 8.5%; Score 247; DB 2; Length 444;
Best Local Similarity 24.4%; Pred. No. 2e-09;
Matches 99; Conservative 69; Mismatches 224; Indels 14; Gaps 9;
QY 87 PGAAQPKPASLGRG-RGAAAAILSLGNVLNLYDRYTVAGVLDDIQHFGVKDRGAGLLQS 145
Db 4 PATSPVIAPIVSTAYRRYALWLVLLIYTLNFLDRQVVNLAEPKEDGLADWQLGMVTG 63
QY 146 VFICSP--MVAIPFYLGDNRKVLSCGIFFFSVNTFFSSFPQPFYLLVLSRLV 203
Db 64 LAFAIYTVLGIPI-ARMAETKNRPIIIGISAVMSAFTVWCGF--AQNFQLILARIGV 120
QY 204 GIGASYSYIAPTIIIGDLFTKNTRLMLSVFYFAIPLGSLGIVTSSVKQAAGDWHAL 263
Db 121 GYGEACTPPANSLISDYVPKRAIAFYIGTIPGLTLAGMAGLVADAYG-WRVAF 179
QY 264 RVSPVLGMITGLILILVPATKRGHADQLGDKARTSWLRDMKALIRNRSYVFSLS 323
Db 180 MVAGAPGLLFALIAFTLVEPRKLAEMAARASTQISFAAALAVLATKTFWLVALAAS 239
QY 324 AVSFATGALGMPIPLYLHRAQ--VVKTAETCNPPCGAKDSLIFGATCTFTGFLGVVTG 381
Db 240 IKAFTGYGYAPTFASFFRVHGPETIAQLAGTFLGSAGFL-GLALGLINGTAGVIGAMLG 298
QY 382 AGATWCRUKTORADPLCAV-GMLGSAIFICLIFVAAKSSIVGAYICIFVGETLFSNW 440
Db 299 GVLAURLGAKDRAVTVVPAISVVTPIPFVAMSLDPAIPMAIGLLSVNALLATLWIGPV 358
QY 441 AITADILMYVITPRATAVALSQSTSHLLGDAGSPYLIGTSLDI 486
Db 359 YATA---QSIQVDPALRATASAVLLLIINLIIGLFGFLVGLLSDL 401

RESULT 13
S48313
hexose transport protein HXT10 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein F008; protein YFL011W
C:Species: Saccharomyces cerevisiae
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C:Accession: S48313; S56243; S62302
R:Churcher, C.
submitted to the EMBL Data Library, September 1994

A:Reference number: S48310
A:Accession: S48313
A:Molecule type: DNA
A:Residues: 1-546 <CHU>
A:Cross-references: UNIPROT:P43581; EMBL:Z46255; NID:G559925; PIDN:CAA86344.1; PID:G55955
R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sasan
submitted to the EMBL Data Library, May 1995
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces c
A:Reference number: S56186
A:Accession: S56243
A:Molecule type: DNA
A:Residues: 1-546 <MUR>
A:Cross-references: EMBL:D50617; NID:G836685; PIDN:BAA09227.1; PID:G836743; MIPS:YFL011
R:Murakami, Y.
submitted to the EMBL Data Library, April 1994
A:Reference number: S62302
A:Accession: S62302
A:Molecule type: DNA
A:Residues: 1-546 <MUR>
A:Cross-references: EMBL:D31600; NID:G836814; PIDN:BAA06493.1; PID:G893435
C:Genetics:
A:Gene: SGD:HXT10
A:Cross-references: SGD:S0001883; MIPS:YFL011W
A:Map position: 6L
C:Superfamily: maltose transport protein MAL61
C:Keywords: sugar transport; transmembrane protein
F:48-64/Domain: transmembrane #status predicted <TM1>
F:101-117/Domain: transmembrane #status predicted <TM2>
F:130-146/Domain: transmembrane #status predicted <TM3>
F:155-171/Domain: transmembrane #status predicted <TM4>
F:224-240/Domain: transmembrane #status predicted <TM5>
F:348-364/Domain: transmembrane #status predicted <TM6>
F:373-389/Domain: transmembrane #status predicted <TM7>
F:412-428/Domain: transmembrane #status predicted <TM8>
F:481-497/Domain: transmembrane #status predicted <TM9>
Query Match 7.5%; Score 219; DB 2; Length 546;
Best Local Similarity 21.7%; Pred. No. 2e-07;
Matches 108; Conservative 84; Mismatches 178; Indels 128; Gaps 24;
QY 107 ILSGNVLNLYDRYTVAGVL--LDIQHFG-----VKDRGAGLLQSVPFCFMVA 155
Db 55 MIAFGGFIQDGTGTISGFINQDTQKRRFEGELQDRGSLSDVTRGLIVGIFNICALGG 114
QY 156 PIFGVLGDRFNRKVLSCGIFFW---SAVTFSSSFIPOQYFWLLVLSGLVIGEASYST 212
Db 115 LTLGRLGDIYGRKIGLMLCVILVYVGVIVIQIASSDKWTQYF-----IGRIVSGMVGGVAV 170
QY 213 IAPTIIIGDLFTKNTRLMLSVFYFAIPLGSLGIVTSSVKQAAGDWHW-----ALR 264
Db 171 LSPFLISISPKHLRGTCVSFYQLMITLIGIFLGICTNYGTKKYSNIOQRVPLGLCFWA 230
QY 265 VSPVLGMITGLILILVPATKRGHADQLGDKARTSWLRDMKALIRNRSYVYF-----SSL 321
Db 231 IFMVIGM-----VMVPEPR-VLVEKGYEARESLAKSNKVTVTPDGVVFEEDTIVA 282
QY 322 TSASVATGALGMPIPLYLHRAQVQKT-----AETCN-----SPPCGAK 361
Db 283 NMELERAVGN-ASWHELFNSKGAIPRVIMGIVIQSLQQLTGCCNYFFYGTTFINAVGHQ 341
QY 362 D-----SLIFGATCTFTGFLGV-VTGAGATWCRUKTKRADPLVCAVGMGSA-IFICLI 415
Db 342 DSFETSIIVLGAVNPASTFVALYIVDKFGRKCLL-----WGSASMAICFVI 387
QY 416 VA-----AKSSIVGAYICIFVGETLIFS---NWAITADILMYTVIPT 454
Db 388 FATVGVTRLWPGKQDPSSQSAGNMIYVTCFFI-----FSFAITWAPIYVIAETPL 442
QY 455 R---RATAVALQSFTSHLLGDAGSPYLIGTSLDIROSTKDSPLWEFLSLGVALMCPV 511
Db 443 VKKRAMAIAVGA--NMWVG-----FLIGFTPTITRSIG-----FSYGIYFMGC--- 485
QY 512 VVLGGMFFLATALPFVSD 529

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 23, 2004, 11:01:18 ; Search time 111 Seconds
(without alignments)
2933.892 Million cell updates/sec

Title: US-10-085-198-48
Perfect score: 2920
Sequence: 1 MMCLCASAAGGAEEDAD.....VHQRGPGGTALAHVRVGAS 566

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_sprot:*

2: uniprot_treml:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|------------|---------------------|
| 1 | 2761 | 94.6 | 619 | 2 Q81VW8 | Q81VW8 homo sapien |
| 2 | 2652 | 90.8 | 590 | 2 Q91VM4 | Q91VM4 mus musculus |
| 3 | 1298 | 44.5 | 528 | 2 Q9H2V7 | Q9H2V7 homo sapien |
| 4 | 1297 | 44.4 | 524 | 2 Q7L541 | Q7L541 homo sapien |
| 5 | 1297 | 44.4 | 528 | 2 Q6FL82 | Q6FL82 homo sapien |
| 6 | 1297 | 44.4 | 528 | 2 AAH65235 | AAH65235 homo sapi |
| 7 | 1289 | 44.1 | 528 | 2 Q8R0G7 | Q8R0G7 mus musculus |
| 8 | 1278 | 43.8 | 528 | 2 Q9EOK0 | Q9EOK0 mus musculus |
| 9 | 1266.5 | 43.4 | 506 | 2 Q8UVB7 | Q8UVB7 brachydanio |
| 10 | 1264.5 | 43.3 | 506 | 2 Q7ZU13 | Q7ZU13 brachydanio |
| 11 | 1258.5 | 43.1 | 486 | 2 Q99LN7 | Q99LN7 mus musculus |
| 12 | 1167 | 40.0 | 476 | 2 Q86VU7 | Q86VU7 homo sapien |
| 13 | 1139 | 39.0 | 514 | 2 Q9D232 | Q9D232 mus musculus |
| 14 | 1115 | 38.2 | 420 | 2 Q71RB5 | Q71RB5 homo sapien |
| 15 | 1115 | 38.2 | 420 | 2 AAQ15259 | AAQ15259 homo sapi |
| 16 | 1088 | 37.3 | 454 | 2 Q8TCS5 | Q8TCS5 homo sapien |
| 17 | 1080 | 37.0 | 512 | 2 Q6ZMD2 | Q6ZMD2 homo sapien |
| 18 | 1080 | 37.0 | 512 | 2 BAD18797 | BAD18797 homo sapi |
| 19 | 1049 | 35.9 | 492 | 2 Q8BYA9 | Q8BYA9 mus musculus |
| 20 | 965 | 33.0 | 643 | 2 Q7Q5N4 | Q7Q5N4 anopheles g |
| 21 | 964.5 | 33.0 | 605 | 2 Q9GQ00 | Q9GQ00 drosophila |
| 22 | 959 | 32.8 | 630 | 2 Q9GQ02 | Q9GQ02 drosophila |
| 23 | 930.5 | 31.9 | 605 | 2 Q9GQ01 | Q9GQ01 drosophila |
| 24 | 925 | 31.7 | 630 | 2 Q9GQ03 | Q9GQ03 drosophila |
| 25 | 916.5 | 31.4 | 477 | 2 Q960X6 | Q960X6 drosophila |
| 26 | 874.5 | 29.9 | 518 | 2 Q9NEW5 | Q9NEW5 caenorhabdi |
| 27 | 855 | 29.3 | 531 | 2 Q01927 | Q01927 caenorhabdi |
| 28 | 816 | 27.9 | 385 | 2 Q81Z31 | Q81Z31 homo sapien |
| 29 | 815 | 27.9 | 488 | 2 Q18545 | Q18545 caenorhabdi |
| 30 | 710 | 24.3 | 401 | 2 Q19235 | Q19235 caenorhabdi |
| 31 | 644 | 22.1 | 229 | 2 Q8N953 | Q8N953 homo sapien |

| | | | | | |
|----|-------|------|-----|------------|---------------------|
| 32 | 633.5 | 21.7 | 402 | 2 Q8MLD0 | Q8MLD0 drosophila |
| 33 | 633.5 | 21.7 | 402 | 2 Q9GQP9 | Q9GQP9 drosophila |
| 34 | 442 | 15.1 | 118 | 2 Q8R119 | Q8R119 mus musculus |
| 35 | 392.5 | 13.4 | 507 | 2 Q9ZQ41 | Q9ZQ41 arabidopsis |
| 36 | 380 | 13.0 | 492 | 2 Q6NMW6 | Q6NMW6 arabidopsis |
| 37 | 380 | 13.0 | 492 | 2 AAS47627 | AAS47627 arabidops |
| 38 | 380 | 13.0 | 492 | 2 AAT41792 | AAT41792 arabidops |
| 39 | 355 | 12.2 | 484 | 2 Q9FLG8 | Q9FLG8 arabidopsis |
| 40 | 354.5 | 12.1 | 746 | 2 Q49546 | Q49546 arabidopsis |
| 41 | 327 | 11.2 | 471 | 2 Q6J2D6 | Q6J2D6 pseudomonas |
| 42 | 327 | 11.2 | 471 | 2 Q88BW0 | Q88BW0 pseudomonas |
| 43 | 327 | 11.2 | 471 | 2 AAT35185 | AAT35185 pseudomon |
| 44 | 324 | 11.1 | 479 | 2 Q87XS1 | Q87XS1 pseudomonas |
| 45 | 319.5 | 10.9 | 431 | 2 Q89KI2 | Q89KI2 bradyrhizob |

ALIGNMENTS

RESULT 1

Q81VW8

ID Q81VW8 PRELIMINARY; PRT; 619 AA.

AC Q81VW8; .

DT 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Similar to spinster-like protein (Fragment).

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Strausberg R.;

RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

DR EMBL; BC041772; AAH41772.1; .

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR007114; MFS.

DR Pfam; PF00083; Sugar tr; 1.

DR PROSITE; PS00850; MFS; 1.

FT Transmembrane.

KW NON_TER

SQ SEQUENCE 619 AA; 64767 MW; 77C223EB281C4286 CRC64;

Query Match 94.6%; Score 2761; DB 2; Length 619;

Best Local Similarity 99.3%; Pred. No. 7, 1e-161;

Matches 535; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MMCLCASAAGGAEEDADERRRRRRRGAQAGAGCGCGAGAGGAGVSAAGDEVQTL 60

Db 71 MMCLCASAAGGAEEDADERRRRRRRGAQAGAGCGCGAGAGGAGVSAAGDEVQTL 130

QY 61 SGSVRRATPGPTGTGCAATAGCAQPKASLGRGCAAAALSLGNVLYLDY 120

Db 131 SGSVRRATPGPTGTGCAATAGCAQPKASLGRGCAAAALSLGNVLYLDY 190

QY 121 TVAGVLLDQQHFGVKDRGAGLLQSVFICSFVVAAPIFGLGDRNRKVLSCGFFFWA 180

Db 191 TVAGVLLDQQHFGVKDRGAGLLQSVFICSFVVAAPIFGLGDRNRKVLSCGFFFWA 250

QY 181 VTFSSSFIPQQYFWLLVLSRGLVGIGEASYSYTIPTIIGDLFTKTRTLMLSVFFFAIPL 240

Db 251 VTFSSSFIPQQYFWLLVLSRGLVGIGEASYSYTIPTIIGDLFTKTRTLMLSVFFFAIPL 310

QY 241 GSGLYITGSSVQKQAGDWHALRVSPVLGMITGTLLILVLPATKRGHADQIGDQKART 300

Db 311 GSGLYITGSSVQKQAGDWHALRVSPVLGMITGTLLILVLPATKRGHADQIGDQKART 370

```
Qy 301 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 360
Dy 371 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 430
Qy 361 KDSLI FGAI TCFTGFLGVVTCAGATRCRLKTKQADPLVCAVGMGLGSAIFCLIFVAAKS 420
Dy 431 KDSLI FGAI TCFTGFLGVVTCAGATRCRLKTKQADPLVCAVGMGLGSAIFCLIFVAAKS 490
Qy 421 SIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVALQSFTHLLGDAGSPYLIG 480
Dy 491 SIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVALQSFTHLLGDAGSPYLIG 550
Qy 481 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHILGE 539
Dy 551 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHILGE 609

RESULT 2
Q91VM4 ID Q91VM4 PRELIMINARY; PRT; 590 AA.
AC Q91VM4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BC011467 protein (Fragment).
GN Name=BC011467;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/fl;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/fl;
RX Strausberg R.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BC011467; AAH11467.1; -
DR MGD; MGI:2384936; BC011467.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
DR Transmembrane.1
FT NON_TER
SQ SEQUENCE 590 AA; 61956 MW; 029412B949E1E0CF CRC64;
```

```
Query Match 90.8%; Score 2652; DB 2; Length 590;
Best Local Similarity 95.4%; Pred. No. 3.2e-154;
Matches 514; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Qy 1 MMCECASAAGAAEEDAEERRRRRGNORAGSGCCGARGAGGAGYSAAGDEVOTL 60
Dy 42 MMCECASAAGAAEEDAEERRRRRGNORAGSGCCGARGAGGAGYSAAGDEVOTL 101
Qy 61 SGSVRRAPTPGPTGTCGCAATAGKPGCAQPKPASLGRGRGAAAAAILSLGNVLNYLDY 120
Dy 102 SGSVRRAPTPGPTGTCGCAATAGKPGCAQPKPASLGRGRGAAAAAILSLGNVLNYLDY 161
Qy 121 TVAGVLLDIQHFQGVKDRGAGLLQSVFICSFVMAAPIFGYLGRDRNRKVLSCGIPFWSA 180
Dy 162 TVAGVLLDIQHFQGVKDRGAGLLQSVFICSFVMAAPIFGYLGRDRNRKVLSCGIPFWSA 221
Qy 181 VTFSSSFIPOQYFWLLVLSRGLVIGGEASYSSTIAPTIIIGDLFTKNTTLMLSVFFYFAIPL 240
Dy 222 VTFSSSFIPOQYFWLLVLSRGLVIGGEASYSSTIAPTIIIGDLFTKNTTLMLSVFFYFAIPL 281
Qy 241 GSGLGVIITGSSVKQAAGDHWALRVSPVLGMITGTLILVLPATKRGHADQLGDLKART 300
Dy 282 GSGLGVIITGSSVKQAAGDHWALRVSPVLGMITGTLILVLPATKRGHADQLGDLKART 341
Qy 301 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 360
Dy 342 SWLRDMKALIRNRSYVFSSLSATSAVSFATGALGMWIPLYLHRAQVQKTAETCNPPCGA 401
Qy 361 KDSLI FGAI TCFTGFLGVVTCAGATRCRLKTKQADPLVCAVGMGLGSAIFCLIFVAAKS 420
Dy 402 KDSLI FGAI TCFTGFLGVVTCAGATRCRLKTKQADPLVCAVGMGLGSAIFCLIFVAAKT 461
Qy 421 SIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVALQSFTHLLGDAGSPYLIG 480
Dy 462 SIVGAYICIFVGETLLFSNWAITADILMYVVIPTRRATAVALQSFTHLLGDAGSPYLIG 521
Qy 481 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHILGE 539
Dy 522 FISDLIRQSTKDSPLWEFLSLGYALMLCPFPVVVLGGMFFLATALFFVSDRARAEOHILGE 580

RESULT 3
Q9H2V7 ID Q9H2V7 PRELIMINARY; PRT; 528 AA.
AC Q9H2V7; Q9BRN5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Spinster (SPINL protein).
GN Name=SPINL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21238517; PubMed=11340170;
RA Nakano Y., Fujitani K., Kurihara J., Ragan J., Ueui-Aoki K.,
RA Shimoda L., Lukacovich T., Suzuki K., Sezaki M., Sano Y., Ueda R.,
RA Awano W., Kaneda M., Umeda M., Yamamoto D.;
RT "Mutations in the novel membrane protein spinster interfere with
RT programmed cell death and cause neural degeneration in Drosophila
RT melanogaster.";
RL Mol. Cell. Biol. 21:3775-3788 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129; TISSUE=Mammary tumor. Brcal-1/fl;
RX Strausberg R.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BC011467; AAH11467.1; -
DR MGD; MGI:2384936; BC011467.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
DR Transmembrane.1
FT NON_TER
SQ SEQUENCE 590 AA; 61956 MW; 029412B949E1E0CF CRC64;
```

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
Blakesley R.W., Goughman J., Schmutz J., Myers R.M., Butterfield Y.S.,
Rodriguez A.C., Grinwood J., Schmutz J., Smailus D.E., Schnerch A., Schein J.E.,
Krzyszynski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
SEQUENCE FROM N.A.
TISSUE=Brain;
Strausberg R.;
Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
TISSUE=Pancreas;
Strausberg R.;
Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC EMBL; AF212371; AAG43830.1; -;
EMBL; BC008325; AAH08325.1; -;
EMBL; BC038961; AAH38961.1; -;
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0005215; P:transporter activity; IEA.
GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 528 AA; 56629 MW; F1B9D2EB3F9F1B48 CRC64;
Query Match 44.5%; Score 1298; DB 2; Length 528;
Best Local Similarity 53.4%; Pred. No. 1.9e-71;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;
QY 53 AGDEVQTLGSGVRRAPTGE-PGTGTPGCAATAGKGAQGPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGFGVGTGPGTGPSTGPKSEEPVQEGQLRITGLSPGSA 61
QY 105 AAILSGNLVNYLDRTYVAGVLLDIOQHFQGVKDRGAGLLQSVFICFMVAAPIFGYLGR 164
DB 62 VAVLCVNLNLYMDRTYVAGVLPDIEQFENIGDSSGLITQTVFISSYMWLAPVFGYLD 121
QY 165 FNRKVLSCGIFPWSAVTSSFPQQYFWLLVSLRGLVIGGEASSTIAPTIIIGDLPTK 224
DB 122 YNRKYLKMGCGIAFWLSVLTGSSFIPEGEHFWLLLTGLVGVGEASSTIAPTIIADLF 181
QY 225 NTRTLMVYFPAIPGSLGIVTGSVKQAGDHWALRVSPVLGMITGLILIVPAT 284
DB 182 DQSRMLSTFPAIPVSGSLGIVTAGSKVDMAGDHWALRVTPGLGVAVLLFLVVRP 241
QY 285 XRGHADQLDQLKAR-TSWLRDMKALIRNSYVFFSLATSAVSFATGALGMWIPYLHRA 343
DB 242 PRGAVERHSDLPNLTNSWADLRALARNPSFVLSGLGFTAVFVTSGLAWAPAFLLRS 301
QY 344 QVYQKTAETC-NSPPCGAKDSLIFGATCTFTGELGVVTCAGATRCWLKATQADPLVCA 402
DB 302 RVVLGETPPLCPGDCSSSDSLIFGLITCLTGLVGLGVGEISRLRHSNPRADPLVCA 361
QY 403 GMLGSAIFCLIFVAKKSIIVGAYICI FVGETLFSNWAITADILMVVYIPTRRTAVAL 462
DB 362 GLGGSAPFLSLACARGIVATYIFIFIGETLLSNMVAIVADILMVVYIPTRRTAEAF 421
QY 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKXDSPLWEFLSLGYALMLCPFFVVLGGMFFLAT 522

Db 422 QIVLSHLGLDAGSPYILIGLISDLRRNWPSPFLSEFRALQFSLMLCAFVGALGGAFLGT 481
QY 523 ALFFVSDRARARBOHLGE--RRAGVR---VHQRG 551
DB 482 AFIETADRRRAQLHQVGLLHEAGSTDDRIIVPQRG 516
RESULT 4
Q7L541
ID Q7L541 PRELIMINARY; PRT; 524 AA.
AC Q7L541;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DE SPINL protein (Fragment).
GN Name=SPINL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins S.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altaeul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.F., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
Blakesley R.W., Goughman J., Schmutz J., Myers R.M., Butterfield Y.S.,
Rodriguez A.C., Grinwood J., Schmutz J., Smailus D.E., Schnerch A., Schein J.E.,
Krzyszynski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
EMBL; BC006156; AAH06156.1; -;
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00850; MFS; 1.
KW Transmembrane.
FT NON_TER
SQ SEQUENCE 524 AA; 56283 MW; CSEBA9DB206256E4 CRC64;
Query Match 44.4%; Score 1297; DB 2; Length 524;
Best Local Similarity 54.8%; Pred. No. 2.2e-71;
Matches 273; Conservative 61; Mismatches 148; Indels 16; Gaps 7;
QY 70 GP-PGTGTPGCAATAGKGAQGPASLGR-----GRGA-AAAILSGNLVNYLDRTY 121
DB 15 GVPVPTGLPGSTGPKSEEPVQEGQLRITGLSPGSAIIVAVLCVNLNLYMDRPT 74
QY 122 VAGVLLDIOQHFQGVKDRGAGLLQSVFICFMVAAPIFGYLGRNKRKVLSCGIFPWSAV 181
DB 75 VAGVLPDIEQFENIGDSSGLITQTVFISSYMWLAPVFGYLDYRNRKVLKMGCIAPVSLV 134
QY 182 TFSFSSFIPOQYFWLLVSLRGLVIGGEASSTIAPTIIIGDLPTKTRTLMVSVYFPAIFLG 241

```
Db 135 TLGSSFIPEHFWLHLLLRGLVGVGEASYSYTIPTLIADLFVADQRSMLSIFYFAIPVG 194
Qy 242 SGLGYITGSSVKQAGDWHMRLVSPVLGMITGTLLILVLPATKRGHADQLGDQKAR-T 300
Db 195 SGLGYIAGSKVDWAGDWHMRLVTPGVGVAVLLFLVVRPRGAVRHSDDLPPNPT 254
Qy 301 SWLRDMKALIRNRSYVFSLSLATSVAFSATGALGMWIPLYLHRAOVQKTAETC-NSPPCG 359
Db 255 SSWADLRALARNPSPVLSLGLTAVAFVTVGSLWAPAFLLRSRVLGETPCLPGDSCS 314
Qy 360 AKDSLIFGAITCTGFLGVTVTGAGATRCWRLKRTORADPLVCAVGMGLSAGIICLIFVAAK 419
Db 315 SDSLIFGLITCLTGLVGLGVGEISRRRLRSHNPRADPLVCAVGMGLSAPFLFLSLACAR 374
Qy 420 SSIIVGAYICIFVGETLLPSNWAITADILMVYVPIPTRRATAVALQSFTHLGDAGSPYLI 479
Db 375 GSIIVATYIFIFIGETLLSMNWAIVADILLYVYVPIPTRRSTABAFQIVLSHLGDAGSPYLI 434
Qy 480 GFISDLIRQSKDPSLWELFSLGVALMLCPFFVVLGGMFFLATALFFVSDRARAEOHLGE 539
Db 435 GLISDRLRNWPSPFLSEFRALQSLMLCAVFGALGGAFLGTAIFIEADRRRAQLHVQ 494
Qy 540 --RRAGVR-----VVHQRG 551
Db 495 LLHEAGSTDDRIWVPQRG 512

RESULT 5
Q6P182
ID Q6P182 PRELIMINARY; PRT; 528 AA.
AC Q6P182;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SPINL protein.
GN Name=SPINL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Prange C.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchwood J.W., Green E., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
DR EMBL; BC065235; LOC65235.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
```

```
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 528 AA; 56655 MW; BC03B6BF2BCB03A6 CRC64;

Query Match 44.4%; Score 1297; DB 2; Length 528;
Best Local Similarity 53.4%; Pred. No. 2.2e-71;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

Qy 53 AGDEVQLSSVRRATGTP-PTGTPCTPGCAATAGKGAQOQKPASLGR-----GRGA-A 104
Db 2 AGSDTAPFLSQADDPDGPVPGTGLPGSTGNPKSEPEVDPQEGQLRITGLSPGSRALI 61
Qy 105 AAILSGNLVNYLDRTVYAGVLLDIOOHFGVKRGAGLLQSVFICSPMAAPIFGYLGD 164
Db 62 VAVLCYNLNYMDRTVYAGVLPDIEQGFNIGDSSGLIQTVFISWVWVLAAPFGYLGDR 121
Qy 165 FNRKVLSCGIFPWSAVTFSSSFIPOQYFWLLVLSRLGVGIGEASYSTIAPTIGDFTK 224
Db 122 YNRKYLKMGCGIAFWSLVTLGSSSFIPEHFWLHLLTRGLVGVGEASYSTIAPTIIADLFVA 181
Qy 225 NTRTLMLSVYFAIPGLSGLYITGSSVKQAGDWHMRLVSPVLGMITGTLLILVLPAT 284
Db 182 DQSRMLSIIFYFAIPVSGGLGYTAGSKVDMAGDWHMRLVTPCLGVVPLLFLVVRREP 241
Qy 285 KRGHADQLGDQKAR-TSWLRDMKALIRNRSYVFSLSLATSVAFSATGALGMWIPLYLHRA 343
Db 242 PRGAVRHSDDLPPNPTSTWADRLARNDPSFVLSLGLTAVAFVTVGSLWAPAFLLRS 301
Qy 344 QVVQKTAETC-NSPPCGAKDSLIFGAITCTGFLGVTVTGAGATRCWRLKRTORADPLVCAV 402
Db 302 RVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRRLRSHNPRADPLVCAV 361
Qy 403 GMLGSAIFICLI FVAAKSSIVGAYICIFVGETLLPSNWAITADILMVYVPIPTRRATAVAL 462
Db 362 GLGSAAPFLFLSLACARGSIATYIFIFIGETLLSMNWAIVADILLYVYVPIPTRRSTAEAF 421
Qy 463 QSFTHSLGDAGSPYLIIGTSLDIROSTKDSPLWELFSLGVALMLCPFFVVLGGMFFLAT 522
Db 422 QIVLSHLGDAGSPYLIIGLISDRLRNWPSPFLSEFRALQSLMLCAVFGALGGAFLGT 481
Qy 523 ALFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
Db 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIWVPQRG 516

RESULT 6
AAH65235 PRELIMINARY; PRT; 528 AA.
AC AAH65235;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE SPINL protein.
GN SPINL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchwood J.W., Green E., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (by similarity).
DR EMBL; BC065235; LOC65235.1; -.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
```



```
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC065235; AAH65235.1; -.
SQ SEQUENCE 528 AA; 56655 MW; BC03B6BF2BCB03A6 CRC64;

Query Match 44.4%; Score 1297; DB 2; Length 528;
Best Local Similarity 53.4%; Pred. No. 2.2e-71;
Matches 275; Conservative 62; Mismatches 162; Indels 16; Gaps 7;

QY 53 AGDEVQTLGSGVRRAPTGP-PGTPGTPGCAATAKPGCAQGPASLGR-----GRGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPFAPGHPGLPGMGNPKSGELEVDPCEGLQRTGLSRGHSTLI 61

QY 105 AAILSLGNVINYLDRTYVAGVLLDIOQHFGVQKDRGAGLLQSVFCSFMVAAPFGYLGDR 164
DB 62 VAVLCYINLNYMDRTVAGVLPDIEQFFNIGDSSGLQTFVFSYMWLAPVFGYLGDR 121

QY 165 FNRKVLSCGIFPWSAVTFSSFPQOYFWLLVSLRGLVIGEASYSTIAPTIGDLFTK 224
DB 122 YNRKYLKMGCGIAPFWSLVTLGSSPIGHEFWLLLTGLVGVGEASYSTIAPTILADLFA 181

QY 225 NRTTLMISVYFAIPGLSGGLGYTGSVQKAGDWHALRVSPVLGMITGTLILVLPAT 284
DB 182 DQSRMLSIIFYFAIPVGSGLGYTAGSKVDKADGWHALRVTPGLGVLLVLPVREP 241

QY 285 KRGHADQLGQPKAR-TSWLRDMKALIRNSYVFFSLATSAYSFATGALGMWIPLYLHRA 343
DB 242 PRGAVEHSDLPPLNTFSWADRLARLNPFSVLSSLGFTAVAFVTGSLALWAPFLRS 301

QY 344 QVQKTAETC-NSPPCGAKDSLIFGAICTFTGFLGVVTVGAGATRCWLKRTQADPLVCV 402
DB 302 RVVLGETPPCLPGDCSSSDSLIFGLITCLTGLVGLGVGEISRRRLHNSPRADPLVCAT 361

QY 403 GMLGSAIFICLIFVAAKSIVGAYICIFVGETLFSNWAITADILMYVVIPTTRATAVAL 462
DB 362 GLUGSAPFLSLACARGSIATYIFIFIGETLLSNMNAIVADILLYVVIPTTRATAEAF 421

QY 463 QGFTSHLLGAGSPYLLIGFISDLIROSTKDSPLWELSLGYALMLCPFVVVLGGMFFLAT 522
DB 422 QIVLSHLLGAGSPYLLIGLISDLRLRNWPPSFSEFALQFSLMLCAFVGLGGAFLQ 481

QY 523 ALFFVSDRARAEOHLGE--RRAGVR---VVRQGR 551
DB 482 AIFIEADRRRAQLHVQGLLHEAGSTDDRIVVVQGR 516

RESULT 7
Q8R0G7 PRELIMINARY; PRT; 528 AA.
AC Q8R0G7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Spinder.
GN Name=2210013K02Rik;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmer C.W., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McKean P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Touchman J.W., Green E.D., Dickson M.C., Butterfield Y.S.,
RA Blakesley R.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BC026854; AAH26854.1; -.
DR MGD; MGI:1920908; 2210013K02Rik.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000566; Lipocin_cytFABP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS00850; MFS; 1.
DR Transmembrane.
SQ SEQUENCE 528 AA; 56708 MW; DDAB448D7D2B9E17 CRC64;

Query Match 44.1%; Score 1289; DB 2; Length 528;
Best Local Similarity 52.5%; Pred. No. 6.8e-71;
Matches 268; Conservative 65; Mismatches 161; Indels 16; Gaps 6;

QY 53 AGDEVQTLGSGVRRAPTGP-PGTPGTPGCAATAKPGCAQGP-----KPASLGRGGA-A 104
DB 2 AGSDTAPFLSQADDDPDGPFAPGHPGLPGMGNPKSGELEVDPCEGLQRTGLSRGHSTLI 61

QY 105 AAILSLGNVINYLDRTYVAGVLLDIOQHFGVQKDRGAGLLQSVFCSFMVAAPFGYLGDR 164
DB 62 VAVLCYINLNYMDRTVAGVLPDIEQFFNIGDSSGLQTFVFSYMWLAPVFGYLGDR 121

QY 165 FNRKVLSCGIFPWSAVTFSSFPQOYFWLLVSLRGLVIGEASYSTIAPTIGDLFTK 224
DB 122 YNRKYLKMGCGIAPFWSLVTLGSSPIGHEFWLLLTGLVGVGEASYSTIAPTILADLFA 181

QY 225 NRTTLMISVYFAIPGLSGGLGYTGSVQKAGDWHALRVSPVLGMITGTLILVLPAT 284
DB 182 DQSRMLSIIFYFAIPVGSGLGYTAGSKVDKADGWHALRVTPGLGVLLVLPVQEP 241

QY 285 KRGHADQ-LGDQLKARTSWLRDMKALIRNSYVFFSLATSAYSFATGALGMWIPLYLHRA 343
DB 242 PRGAVEHSGSPPLSPSTSWADRLARLNPFSVLSSLGFTSVAFVTGSLALWAPFLRS 301

QY 344 QVQKTAETC-NSPPCGAKDSLIFGAICTFTGFLGVVTVGAGATRCWLKRTQADPLVCV 402
DB 302 RVVLGETPPCLPGDCSSSDSLIFGLITCLTGLVGLGVGEISRRRLHNSPRADPLVCV 361

QY 403 GMLGSAIFICLIFVAAKSIVGAYICIFVGETLFSNWAITADILMYVVIPTTRATAVAL 462
DB 362 GLUGSAPFLSLACARGSIATYIFIFIGETLLSNMNAIVADILLYVVIPTTRATAEAF 421
```

```

Db 362 GLGSAFFLALACARGSIATVIFIFIGETLLSMNWAIVADILLYVVIPTREASTAEAF 421
Qy 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWELFSLGLVALMCLCPVVVLGGMFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGLISDLRRSRWPPPSLSEFRALQFSLMLCAFVGLGAAFLGT 481
Qy 523 ALFFVSDRARAEOHLGERRAGVRVVRHQRGP 552
Db 482 AMFIEDDRRAQLHV-----QGLLHESGP 505

RESULT 8
Q9EQK0 PRELIMINARY; PRT; 528 AA.
ID Q9EQK0
AC Q9EQK0
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Spinster-like protein.
GN Name=2210013K02R1k;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21238517; PubMed=11340170;
RA Nakano Y., Fujitani K., Kurihara J., Ragan J., Usui-Aoki K.,
RA Shimoda L., Lukacsovich T., Suzuki K., Sezaki M., Sano Y., Ueda R.,
RA Awano M., Kaneda M., Umeda M., Yamamoto D.;
RT "Mutations in the novel membrane protein spinster interfere with
RT programmed cell death and cause neural degeneration in Drosophila
RT melanogaster.";
RL Mol. Cell. Biol. 21:3775-3788(2001).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AF212372; AAG43831.1; -
DR MGD; MGI:1920508; 2210013K02R1k.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005566; Lipocin_cytFABP.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS00213; LIFOCALIN; UNKNOWN_1.
DR PROSITE; PS50850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 528 AA; 56860 MW; 85D4029B9B69827C CRC64;

Query Match 43.8%; Score 1278; DB 2; Length 528;
Best Local Similarity 52.2%; Pred. No. 3.2e-70;
Matches 266; Conservative 64; Mismatches 164; Indels 16; Gaps 6;

Qy 53 AGDEVQTLGSGVRRAPTPGP-FPTGTPGCAATAKPGGAQGP-----KPASLGRGRGA-A 104
Db 2 AGSDTAPFLSQADDDPGAPGHPGLPGPMGNPKSGELEVPDCEGLQRIITGLSRGHSITLI 61

Qy 105 AAILSLGNVLYLDRTYVAGVLLDITQHPGVKDRGAGLLQSVFICSFVWAAPIFGYLGDR 164
Db 62 VVVLICYINLLNMDRFTVAGVLTDEQFNIGDGTGLIQTIVFTISSYMLVAPVFGYLGDR 121

Qy 165 FNKKVILSCGIFFWASVTPSSSFIPOQYFWLLVLSRGLVIGGEASYSYTIPTIIGDLPTK 224
Db 122 YNRKFMCGGTAFWSLVLTGSSSFIPEHFVLLFLTRGMVGVGEASYSYTIPTIADFPVA 181

Qy 225 NRTTLMLSVFFPAIPGLSGGLGYITGSSVVKQAGDWHWALRVSPVLGMITGTLILVPA 284
Db 182 DQSRMLSIFFPAIPVSGSLGYIAGSKVKDVAGDWHWALRVTPGLGLVALVLLFLVWQEP 241

Qy 285 KGHADQ-LGDLQKARTSWLRDMKALINRSYVFSLSLATSVAVSPATGALGWIPLYLHRA 343
Db 242 PRGAVERHSGSPPLSPSTSWADLKALARNPSPVLSLGLFTSVAFVFTGSLMALWAPFLRS 301

```

```

Qy 344 QVYQKTAETC-NSPPCGAKDSLIFGALTCTGFLGVVVTGAGATRCWLKTKQRADPLVCAV 402
Db 302 RVVLGETPPCLPGDCSSDSLIFGLITCLTGVGLGMEISRRLRRFNPRADPLVCAA 361
Qy 403 GMLGSAIFICLIIVAAGKSIVGAYICIFVGETLLFNSWAILMVMVVIPTTRATAVAL 462
Db 362 GLGSAFFLALACARGSIATVIFIFIGETLLSMNWAIVADILLYVVIPTREASTAEAF 421
Qy 463 QSFTSHLLGDAGSPYLIGFISDLIRQSTKDSPLWELFSLGLVALMCLCPVVVLGGMFLAT 522
Db 422 QIVLSHLLGDAGSPYLIGLISDLRRSRWPPPSLSEFRALQFSLMLCAFVGLGAAFLGT 481
Qy 523 ALFFVSDRARAEOHLGERRAGVRVVRHQRGP 552
Db 482 AMFIEDDRRAQLHV-----QGLLHESGP 505

RESULT 9
Q8UVB7 PRELIMINARY; PRT; 506 AA.
ID Q8UVB7
AC Q8UVB7
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Not really started.
GN Name=spinl; Synonym=nsrs;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21826209; PubMed=11836794;
RA Young R.M., Marty S., Nakano Y., Wang H., Yamamoto D., Lin S.,
RA Allende M.L.;
RT "Zebrafish yolk-specific not really started (nsr) gene is a vertebrate
RT homolog of the Drosophila spinster gene and is essential for
RT embryogenesis.";
RL Dev. Dyn. 223:298-305(2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AF465772; AAL69987.1; -
DR ZFIN; ZDB-GENE-020228-1; spinl.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PROSITE; PS50850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 506 AA; 54962 MW; 033B0DE912CC1ED2 CRC64;

```

```

Query Match 43.4%; Score 1266.5; DB 2; Length 506;
Best Local Similarity 56.6%; Pred. No. 1.6e-63;
Matches 252; Conservative 66; Mismatches 124; Indels 3; Gaps 2;

Qy 107 ILSLGNVLYLDRTYVAGVLLDITQHPGVKDRGAGLLQSVFICSFVWAAPIFGYLGDRFN 166
Db 53 VLYCYINLLNMDRFTVAGVLTDEHFFGIGDGTGSLQTVFICSMFLAPLFGYLGDRYN 112

Qy 167 RKVLISCGIFFWASVTPSSSFIPOQYFWLLVLSRGLVIGGEASYSYTIPTIIGDLTKNT 226
Db 113 RKLIMCVGIFFWASVTVLASSFIKDFHWALLTRGLVGVGEASYSYTIPTIADLFVKEK 172

Qy 227 RTTLMLSVFFPAIPGLSGGLGYITGSSVVKQAGDWHWALRVSPVLGMITGTLILVPA 286
Db 173 RTNMLSIFFPAIPVSGSGYIVGSKVDVAKDWHWALRVTPGLGLVALVLLFLVWQEP 232

Qy 287 GHADQLGDLQKARTSWLRDMKALINRSYVFSLSLATSVAVSPATGALGWIPLYLHRAQV 346
Db 233 GAIEAHPEHTLHRTSWLADMKALCRNPSPFIIISTPGFTAVAFVFTGSLMALWAPFLRAGV 292

```

```

QY 347 QKTAETCNSPPCGAKDSLIFGAICTCTGFLGVVGTGATRCWCLKTQRADPLVCVAGMLG 406
Db 293 TGVKQPCFKAPCDDSDSLIFGAICTCTGFLGVVGTGATRCWCLKTQRADPLVCVAGMLG 352
QY 407 SAIFICLIFVAAKSSIVGAYICIFVGETLLFNSWALTADILMYWIPTRATRAVALQSF 466
Db 353 AAPFLYLSIFAQASTVATYVIFLGETFLSMNMAIVADILLVWIPTRSTAEAFQIVL 412
QY 467 SHLLGDAGSPYLLIGFISDLIRQSTKDSPLWEFLSLGVALMCLPFVVVLGGMFFLATALPF 526
Db 413 SHLLGDAISPYLLIGVVSIDSIKES--NSYMWEPFSLQMSLLCSFVAVAGAPFLATAVEI 470
QY 527 VSDRABAEQHLGERRAGVRVHVHQRG 551
Db 471 EKDRDLAENYVPSDDAPI-VVPRSG 494

RESULT 10
ID Q7ZUI3 PRELIMINARY; PRT; 506 AA.
AC Q7ZUI3;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Spintester-like.
GN Names=spini;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RA Strausberg R.;
RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
EMBL; BC048024; AAH48024.1; -.
DR ZFIN; ZDB-GENE-020228-1; spin1.
DR GO; GO:0015021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transporter; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR Pfam; PF00083; Sugar tr.; 1.
DR PROSITE; PS50850; MFS; 1.
KW Transmembrane.
SQ SEQUENCE 506 AA; 54980 MW; B416E84D2CC19BF CRC64;

```

```

Query Match 43.3%; Score 1264.5; DB 2; Length 506;
Best Local Similarity 56.6%; Pred. No. 2.1e-69;
Matches 252; Conservative 66; Mismatches 124; Indels 3; Gaps 2;

QY 107 ILSGLNVLNDRYTVAGVLLDIQOHFGVKGDRGAGLLQSVFICSFWAAPIFGVLGDRFN 166
Db 53 VLICYNLLNMDRFTVAGVLPDIHFFGIGDGTGSLLOTVPICSYMFLAPLFGVLGDRYN 112
QY 167 RKVILSCGIFPWSAVTFSSFPQOYFWLLVLRSGLVIGEGEASYSYTIPTIIGDLFTKNT 256
Db 113 RKLIMCVGIFPWSVTVLASSPFGKDFHFWALLTRGLVGVGEGEASYSYTIPTIADLFVKEK 172
QY 227 RTMLSLVFYFAIPLGSLGYITGSSVYKQAGDHWALRVSPVLGMITGTLLILVPAIKR 286
Db 173 RTNMLISIFYFAIPVSGMGYIVGSKVDYVAKDHWALRVTPGLGLLAVELLMLVQEPKR 232
QY 287 GHADQLGDOQLKARTSWLRDMKALINRSYVFSLSATSVAIPATGALGMIPLYLHRAQV 346
Db 233 GAIEAHPETHLRTLWADMKALCRNPSEILSTFGTFAVAVFTVGSALWAPAPLFRAGVP 292
QY 347 QKTAETCNSPPCGAKDSLIFGAICTCTGFLGVVGTGATRCWCLKTQRADPLVCVAGMLG 406
Db 293 TGVKQPCFKAPCDDSDSLIFGAICTCTGFLGVVGTGATRCWCLKTQRADPLVCVAGMLG 352
QY 407 SAIFICLIFVAAKSSIVGAYICIFVGETLLFNSWALTADILMYWIPTRATRAVALQSF 466
Db 353 AAPFLYLSIFAQASTVATYVIFLGETFLSMNMAIVADILLVWIPTRSTAEAFQIVL 412
QY 467 SHLLGDAGSPYLLIGFISDLIRQSTKDSPLWEFLSLGVALMCLPFVVVLGGMFFLATALPF 526
Db 413 SHLLGDAISPYLLIGVVSIDSIKES--NSYMWEPFSLQMSLLCSFVAVAGAPFLATAVEI 470
QY 527 VSDRABAEQHLGERRAGVRVHVHQRG 551
Db 471 EKDRDLAENYVPSDDAPI-VVPRSG 494

RESULT 11
Q99LN7 PRELIMINARY; PRT; 486 AA.
ID Q99LN7;
AC Q99LN7;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE 2210013K02Bik protein (fragment).
GN Name=2210013K02Bik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH 11;
RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Richards S., Worley K.C., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitling M., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human

```



```

Q71RB5
ID Q71RB5; PRELIMINARY; PRT; 420 AA.
AC AAQ15259;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE PP2030.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y.; Zhou X.M.; Zhang P.P.; Jiang H.Q.; Qin W.X.; Zhao X.T.;
RA Wan D.F.; Gu J.R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF370423; AAQ15259.1; -.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS50850; MFS; 1.
SQ SEQUENCE 420 AA; 45301 MW; 5771BC5C252CD048 CRC64;

Query Match 38.2%; Score 1115; DB 2; Length 420;
Best Local Similarity 56.6%; Pred. No. 2.4e-60;
Matches 231; Conservative 50; Mismatches 119; Indels 8; Gaps 4;

QY 152 MVAAPIFGYLGDPRNKRKVLSCGIFFWSAVTFSSSFIPOQYFWLILVLSRGLVGIGEASYS 211
Db 1 MVLAPVGYLGDPRNKRKVLSCGIFFWSAVTFSSSFIPOQYFWLILVLSRGLVGIGEASYS 60

QY 212 TIAPTIIGDLFTKNTRTLMLSVFYFAIPGLGSLGYITGSSVKQAAGDMHWALRVSPVLGM 271
Db 61 TIAPTLIADLFVADQSRMLSIIFYFAIPVSGSLGYIAGSKVMDMAGDMHWALRVTPGLGV 120

QY 272 ITGTLILVLPATKRGHADQDQDKAR-TSWLRDMKALIRNSYVFSLSLATSVAVSPTG 330
Db 121 VAVLLFLVVRPPRGAVRHSDLPPLNPTSMWADLRALARNPSFVLSLGLFTAVAPVTG 180

QY 331 ALGWMIPLYLHRAQVVKTAETC-NSPPCGAKDSLIFGAICTCFTGFLGVVGTGATRWCR 389
Db 181 SLALWAPAFLLRSRVVLGETPPCLPGDSCSSDSLIFGLITCLTGLVGLGVGEISRRLR 240

QY 390 LKQORADPLVCAGVGLGSAIFICLIFVAAKSSIVGAYICIFVGETLLFSNWAITADILMY 449
Db 241 HSNPRADPLVCATGLLGSAPFLFLSLACARGSIVATYIFIFIGETLLSMNWAIVADILLY 300

QY 450 VVIPTRRATAVALQSFTSHLGDAGSPYLIGFISDLIRQSTKDSPLWEFLSLGYALMLCP 509
Db 301 VVIPTRRATAEQIVLSHLGLDAGSPYLIGLISDLRLRNWPPSFLSEFRALQFSLMLCA 360

QY 510 FVVVLGGMFFLATLFFVSDRARAEOHLGE--RRAGVR-----VVHQRG 551
Db 361 FVGALGGAAPFLGTAIFTEADRRRAQLHVQGLLHEAGSTDDRIVVPPQRG 408

Search completed: October 23, 2004, 13:14:07
Job time : 115 secs

```

```

RESULT 15
AAQ15259 PRELIMINARY; PRT; 420 AA.
ID AAQ15259;
AC AAQ15259;
DT 02-MAR-2004 (TREMBlrel. 27, Created)
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
DE PP2030.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang Y.; Zhou X.M.; Zhang P.P.; Jiang H.Q.; Qin W.X.; Zhao X.T.;
RA Wan D.F.; Gu J.R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

```